

## THE NATIONAL ACADEMIES

### *Advisers to the Nation on Science, Engineering, and Medicine*

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

[www.national-academies.org](http://www.national-academies.org)

## Contents

<b>SUMMARY .....</b>	<b>3</b>
<b>1 INTRODUCTION.....</b>	<b>16</b>
Formaldehyde and the Draft Assessment, 16	
The Committee's Task and Approach, 20	
Organization of Report, 22	
References, 23	
<b>2 REVIEW OF METHODS .....</b>	<b>24</b>
Review of the Methodology of the Draft IRIS Assessment	
of Formaldehyde, 25	
Summary, 28	
References, 28	
<b>3 TOXICOKINETICS AND MODES OF ACTION</b>	
<b>OF FORMALDEHYDE.....</b>	<b>29</b>
Toxicokinetics, 30	
Carcinogenesis: Has a Mode of Action of Formaldehyde	
Been Identified?, 44	
Use of a Biologically Based Dose-Response Model, 46	
Conclusions and Recommendations, 58	
References, 60	
<b>4 PORTAL-OF-ENTRY HEALTH EFFECTS .....</b>	<b>64</b>
Irritation, 65	
Decreased Pulmonary Function, 71	
Noncancer Respiratory Tract Pathology, 74	
Asthma, 78	
Respiratory Tract Cancers, 83	
References, 88	

5	<b>SYSTEMIC HEALTH EFFECTS</b> .....	92
	Immunotoxicity, 93	
	Neurotoxicity, 97	
	Reproductive and Developmental Toxicity, 102	
	Lymphohematopoietic Cancers, 108	
	References, 114	
6	<b>REFERENCE CONCENTRATIONS FOR NONCANCER EFFECTS AND UNIT RISKS FOR CANCER</b> .....	118
	Formaldehyde Reference Concentrations, 119	
	Formaldehyde Unit Risks For Cancer, 133	
	Conclusions and Recommendations, 144	
	References, 146	
7	<b>A ROADMAP FOR REVISION</b> .....	151
	Critical Revisions of the Current Draft IRIS Assessment of Formaldehyde, 151	
	Future Assessments and the IRIS Process, 152	
	References, 166	

#### APPENDIXES

A	<b>BIOGRAPHIC INFORMATION ON THE COMMITTEE TO REVIEW EPA'S DRAFT IRIS ASSESSMENT OF FORMALDEHYDE</b> .....	168
B	<b>WEIGHT-OF-EVIDENCE DESCRIPTIONS FROM U.S. ENVIRONMENTAL PROTECTION AGENCY GUIDELINES</b> .....	174

#### BOXES, FIGURES, AND TABLES

##### BOXES

1-1	Statement of Task, 20
-----	-----------------------

##### FIGURES

S-1	Illustration of potential process for identifying an RfC, 13
1-1	Formaldehyde chemical structure, 17
1-2	Formaldehyde concentration in various environments, 18
1-3	Timeline of the development of the draft IRIS assessment, 19
2-1	Elements of the IRIS process, 25
3-1	Schematic representation of the mammalian nasal epithelium, 32

## *Contents*

*xiii*

- 4-1 Odds ratios for physician-diagnosed asthma in children associated with in-home formaldehyde concentrations in air, 82
- 5-1 Origins of lymphohematopoietic cancers, 109
- 5-2 Relative incidence and estimated annual new diagnoses of common lymphohematopoietic cancer subtypes in the United States, 110
- 6-1 Illustration of EPA's process for deriving a reference concentration for formaldehyde, 120
- 6-2 Illustration of a potential process for identifying an RfC from a full database, 132
- 7-1 New IRIS assessment process, 154
- 7-2 Elements of the key steps in the development of a draft IRIS assessment, 155
- 7-3 Example of an article-selection process, 159

## **TABLES**

- 3-1 Analysis of 3D CFD Models by Kimbell et al. (2001a,b) and Overton et al. (2001) for Rat, Monkey, and Human Airways, 41
- 3-2 Overview of the Conolly et al. BBDR Models, 48
- 3-3 Effects of Different Parameters on Predicted Results of the Conolly et al. BBDR Models, 52
- 6-1 Derivation of Candidate RfCs by EPA, 122
- 6-2 Cancer Unit Risk Estimates for Formaldehyde, 142
- 7-1 Criteria for Determining Causality, 157
- 7-2 Hierarchy for Classifying Strength of Causal Inferences on the Basis of Available Evidence, 157

formaldehyde exposure and the three kinds of cancer, EPA's decision to calculate unit risk values for them appears to be defensible on the basis of the agency's cancer guidelines. However, EPA should provide a clear description of the criteria that it used to select the specific cancers and demonstrate a systematic application of the criteria. The calculation of the unit risk values is a complex process, involves many sources of uncertainty and variability, and is influenced by the low-dose extrapolation used (for example, linear vs threshold). The committee therefore recommends that EPA conduct an independent analysis of the dose-response models to confirm the degree to which the models fit the data appropriately. EPA is encouraged to consider the use of alternative extrapolation models for the analysis of the cancer data; this is especially important given the use of a single study, the inconsistencies in the exposure measures, and the uncertainties associated with the selected cancers.

#### **THE FORMALDEHYDE IRIS ASSESSMENT: THE PATH FORWARD**

The committee recognizes that the completion of the formaldehyde IRIS assessment is awaited by diverse stakeholders, and it has tried to be judicious in its recommendations of specific changes noted in its report. However, the committee concludes that the following general recommendations are critical to address in the revision of the draft assessment. First, rigorous editing is needed to reduce the volume of the text substantially and address the redundancies and inconsistencies; reducing the text could greatly enhance the clarity of the document. Second, Chapter 1 of the draft assessment needs to discuss more fully the methods of the assessment. The committee is recommending not the addition of long descriptions of EPA guidelines but rather clear concise statements of criteria used to exclude, include, and advance studies for derivation of the RfCs and unit risk estimates. Third, standardized evidence tables that provide the methods and results of each study are needed for all health outcomes; if appropriate tables were used, long descriptions of the studies could be moved to an appendix or deleted. Fourth, all critical studies need to be thoroughly evaluated for strengths and weaknesses by using uniform approaches; the findings of these evaluations could be summarized in tables to ensure transparency. Fifth, the rationales for selection of studies that are used to calculate RfCs and unit risks need to be articulated clearly. Sixth, the weight-of-evidence descriptions need to indicate the various determinants of "weight." The reader needs to be able to understand what elements (such as consistency) were emphasized in synthesizing the evidence.

The committee is concerned about the persistence of problems encountered with IRIS assessments over the years, especially given the multiple groups that have highlighted them, and encourages EPA to address the problems with development of the draft assessments that have been identified. The committee recognizes that revision of the approach will involve an extensive effort by EPA staff and others, and it is not recommending that EPA delay the revision of the

*Summary*

15

formaldehyde assessment to implement a new approach. However, models for conducting IRIS assessments more effectively and efficiently are available, and the committee provides several examples in the present report. Thus, EPA might be able to make changes in its process relatively quickly by selecting and adapting existing approaches. As exemplified by the recent revision of the approach used for the National Ambient Air Quality Standards, this task is not insurmountable. If the methodologic issues are not addressed, future assessments may still have the same general and avoidable problems that are highlighted here.

## A Roadmap for Revision

In reviewing the draft assessment *Toxicological Review of Formaldehyde-Inhalation Assessment: In Support of Summary Information on the Integrated Risk Information System (IRIS)*, the committee initially evaluated the general methodology (Chapter 2) and then considered the dosimetry and toxicology of formaldehyde (Chapter 3) and the review of the evidence and selection of studies related to noncancer and cancer outcomes (Chapters 4 and 5). Finally, the committee addressed the calculation of the reference concentrations (RfCs) for noncancer effects and the unit risks for cancer and the treatment of uncertainty and variability (Chapter 6). In this chapter, the committee provides general recommendations for changes that are needed to bring the draft to closure. On the basis of “lessons learned” from the formaldehyde assessment, the committee offers some suggestions for improvements in the IRIS development process that might help the Environmental Protection Agency (EPA) if it decides to modify the process. As noted in Chapter 2, the committee distinguishes between the process used to generate the draft IRIS assessment (that is, the development process) and the overall process that includes the multiple layers of review. The committee is focused on the development of the draft IRIS assessment.

### CRITICAL REVISIONS OF THE CURRENT DRAFT IRIS ASSESSMENT OF FORMALDEHYDE

The formaldehyde draft IRIS assessment has been under development for more than a decade (see Chapter 1, Figure 1-3), and its completion is awaited by diverse stakeholders. Here, the committee offers general recommendations—in addition to its specific recommendations in Chapters 3-6—for the revisions that are most critical for bringing the document to closure. Although the committee suggests addressing some of the fundamental aspects of the approach to generating the draft assessment later in this chapter, it is not recommending that the assessment for formaldehyde await the possible development of a revised ap-

proach. The following recommendations are viewed as critical overall changes needed to complete the draft IRIS assessment:

- To enhance the clarity of the document, the draft IRIS assessment needs rigorous editing to reduce the volume of text substantially and address redundancy and inconsistency. Long descriptions of particular studies, for example, should be replaced with informative evidence tables. When study details are appropriate, they could be provided in appendixes.
- Chapter 1 needs to be expanded to describe more fully the methods of the assessment, including a description of search strategies used to identify studies with the exclusion and inclusion criteria clearly articulated and a better description of the outcomes of the searches (a model for displaying the results of literature searches is provided later in this chapter) and clear descriptions of the weight-of-evidence approaches used for the various noncancer outcomes. The committee emphasizes that it is not recommending the addition of long descriptions of EPA guidelines to the introduction, but rather clear concise statements of criteria used to exclude, include, and advance studies for derivation of the RfCs and unit risk estimates.
- Standardized evidence tables for all health outcomes need to be developed. If there were appropriate tables, long text descriptions of studies could be moved to an appendix or deleted.
- All critical studies need to be thoroughly evaluated with standardized approaches that are clearly formulated and based on the type of research, for example, observational epidemiologic or animal bioassays. The findings of the reviews might be presented in tables to ensure transparency. The present chapter provides general guidance on approaches to reviewing the critical types of evidence.
- The rationales for the selection of the studies that are advanced for consideration in calculating the RfCs and unit risks need to be expanded. All candidate RfCs should be evaluated together with the aid of graphic displays that incorporate selected information on attributes relevant to the database.
- Strengthened, more integrative, and more transparent discussions of weight of evidence are needed. The discussions would benefit from more rigorous and systematic coverage of the various determinants of weight of evidence, such as consistency.

#### **FUTURE ASSESSMENTS AND THE IRIS PROCESS**

This committee's review of the draft IRIS assessment of formaldehyde identified both specific and general limitations of the document that need to be addressed through revision. The persistence of limitations of the IRIS assessment methods and reports is of concern, particularly in light of the continued evolution of risk-assessment methods and the growing societal and legislative pressure to evaluate many more chemicals in an expedient manner. Multiple

groups have recently voiced suggestions for improving the process. The seminal "Red Book," the National Research Council (NRC) report *Risk Assessment in the Federal Government: Managing the Process*, was published in 1983 (NRC 1983). That report provided the still-used four-element framework for risk assessment: hazard identification, dose-response assessment, exposure assessment, and risk characterization. Most recently, in the "Silver Book," *Science and Decisions: Advancing Risk Assessment*, an NRC committee extended the framework of the Red Book in an effort to make risk assessments more useful for decision-making (NRC 2009). Those and other reports have consistently highlighted the necessity for comprehensive assessment of evidence and characterization of uncertainty and variability, and the Silver Book emphasizes assessment of uncertainty and variability appropriate to the decision to be made.

*Science and Decisions: Advancing Risk Assessment* made several recommendations directly relevant to developing IRIS assessments, including the draft formaldehyde assessment. First, it called for the development of guidance related to the handling of uncertainty and variability, that is, clear definitions and methods. Second, it urged a unified dose-response assessment framework for chemicals that would link understanding of disease processes, modes of action, and human heterogeneity among cancer and noncancer outcomes. Thus, it suggested an expansion of cancer dose-response assessments to reflect variability and uncertainty more fully and for noncancer dose-response assessments to reflect analysis of the probability of adverse responses at particular exposures. Although that is an ambitious undertaking, steps toward a unifying framework would benefit future IRIS assessments. Third, the Silver Book recommended that EPA assess its capacity for risk assessment and take steps to ensure that it is able to carry out its challenging risk-assessment agenda. For some IRIS assessments, EPA appears to have difficulty in assembling the needed multidisciplinary teams.

The committee recognizes that EPA has initiated a plan to revise the overall IRIS process and issued a memorandum that provided a brief description of the steps (EPA 2009a). Figure 7-1 illustrates the steps outlined in that memorandum. The committee is concerned that little information is provided on what it sees as the most critical step, that is, completion of a draft IRIS assessment. In the flow diagram, six steps are devoted to the review process, and thus the focus of the revision appears to be on the steps after the assessment has been generated. Although EPA may be revising its approaches for completing the draft assessment (Step 1 in Figure 7-1), the committee could not locate any other information on the revision of the IRIS process. Therefore, the committee offers some suggestions on the development process.

In providing guidance on revisions of the IRIS development process (that is, Step 1 as illustrated in Figure 7-1), the committee begins with a discussion of the current state of science regarding reviews of evidence and cites several examples that provide potential models for IRIS assessments. The

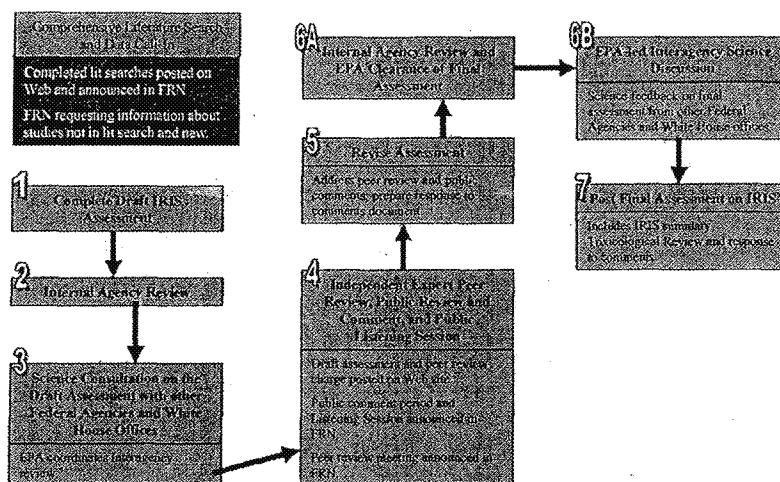


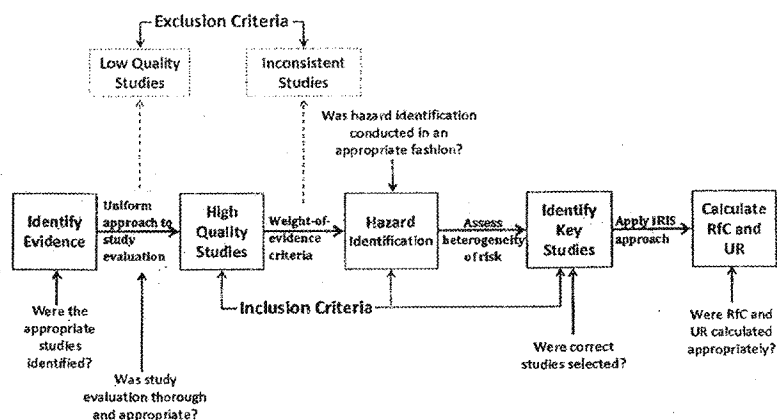
FIGURE 7-1 New IRIS assessment process. Abbreviations: FRN, Federal Register Notice; IRIS, Integrated Risk Information System; and EPA, Environmental Protection Agency. Source: EPA 2009a.

committee also describes the approach now followed in reviewing and synthesizing evidence related to the National Ambient Air Quality Standards (NAAQSs), a process that has been modified over the last 2 years. It is provided as an informative example of how the agency was able to revise an entrenched process in a relatively short time, not as an example of a specific process that should be adopted for the IRIS process. Finally, the committee offers some suggestions for improving the IRIS development process, providing a "roadmap" of the specific items for consideration.

### An Overview of the Development of the Draft IRIS Assessment

In Chapter 2, the committee provided its own diagram (Figure 2-1) describing the steps used to generate the draft IRIS assessment. For the purpose of offering committee comments on ways to improve those steps, that figure has been expanded to indicate the key outcomes at each step (Figure 7-2). For each of the steps, the figure identifies the key questions addressed in the process. At the broadest level, the steps include systematic review of evidence, hazard identification using a weight-of-evidence approach, and dose-response assessment.

The systematic review process is undertaken to identify all relevant literature on the agent of interest, to evaluate the identified studies, and possibly to



**FIGURE 7-2** Elements of the key steps in the development of a draft IRIS assessment. Abbreviations: IRIS, Integrated Risk Information System; RfC, reference concentration; and UR, unit risk.

provide a qualitative or quantitative synthesis of the literature. Chapter 1 of the draft IRIS assessment of formaldehyde provides a brief general description of the process followed by EPA, including the approach to searching the literature. However, neither Chapter 1 nor other chapters of the draft provide a sufficiently detailed description of the approach taken in evaluating individual studies. In discussing particular epidemiologic studies, a systematic approach to study evaluation is not provided. Consequently, some of the key methodologic points are inconsistently mentioned, such as information bias and confounding.

For hazard identification, the general guidance is also found in Chapter 1 of the draft IRIS assessment. The approach to conducting hazard identification is critical for the integrity of the IRIS process. The various guidelines cited in Chapter 1 provide a general indication of the approach to be taken to hazard identification but do not offer a clear template for carrying it out. For the formaldehyde assessment, hazard identification is particularly challenging because the outcomes include cancer and multiple noncancer outcomes. The various EPA guidelines themselves have not been harmonized, and they provide only general guidance. Ultimately, the quality of the studies reviewed and the strength of evidence provided by the studies for deriving RfCs and unit risks need to be clearly presented. More formulaic approaches are followed for calculation of RfCs and unit risks. The key issue is whether the calculations were conducted appropriately and according to accepted assessment procedures.

### Brief Review of Established Best Practices

The following sections highlight some best practices of current approaches to evidence-based reviews, hazard identification, and dose-response assessment that could provide EPA guidance if it decides to address some of the fundamental issues identified by the committee. The discussion is meant not to be comprehensive or to provide all perspectives on the topics but simply to highlight some important aspects of the approaches. The committee recognizes that some of the concepts and approaches discussed below are elementary and are addressed in some of EPA's guidelines. However, the current state of the formaldehyde draft IRIS assessment suggests that there might be a problem with the practical implementation of the guidelines in completing the IRIS assessments. Therefore, the committee highlights aspects that it finds most critical.

### Current Approaches to Evidence-Based Reviews

Public-health decision-making has a long history of using comprehensive reviews as the foundation for evaluating evidence and selecting policy options. The landmark 1964 report of the U.S. surgeon general on tobacco and disease is exemplary (DHEW 1964). It used a transparent method that involved a critical survey of all relevant literature by a neutral panel of experts and an explicit framework for assessing the strength of evidence for causation that was equivalent to hazard identification (Table 7-1).

The tradition of comprehensive, evidence-based reviews has been continued in the surgeon general's reports. The 2004 surgeon general's report, which marked the 40th anniversary of the first report, highlighted the approach for causal inference used in previous reports and provided an updated and standardized four-level system for describing strength of evidence (DHHS 2004) (Table 7-2).

The same systematic approaches have become fundamental in many fields of clinical medicine and public health. The paradigm of "evidence-based medicine" involves the systematic review of evidence as the basis of guidelines. The international Cochrane Collaboration engages thousands of researchers and clinicians throughout the world to carry out reviews. In the United States, the Agency for Healthcare Research and Quality supports 14 evidence-based practice centers to conduct reviews related to healthcare.

There are also numerous reports from NRC committees and the Institute of Medicine (IOM) that exemplify the use of systematic reviews in evaluating evidence. Examples include reviews of the possible adverse responses associated with Agent Orange, vaccines, asbestos, arsenic in drinking water, and secondhand smoke. A 2008 IOM report, *Improving the Presumptive Disability Decision-Making Process for Veterans*, proposed a comprehensive new scheme for

**TABLE 7-1 Criteria for Determining Causality**

Criterion	Definition
Consistency	Persistent association among different studies in different populations
Strength of association	Magnitude of the association
Specificity	Linkage of specific exposure to specific outcome
Temporality	Exposure comes before effect
Coherence, plausibility, analogy	Coherence of the various lines of evidence with a causal relationship
Biologic gradient	Presence of increasing effect with increasing exposure (dose-response relationship)
Experiment	Observations from "natural experiments," such as cessation of exposure (for example, quitting smoking)

Source: DHHS 2004.

**TABLE 7-2 Hierarchy for Classifying Strength of Causal Inferences on the Basis of Available Evidence**

- A. Evidence is *sufficient* to infer a causal relationship.
- B. Evidence is *suggestive but not sufficient* to infer a causal relationship.
- C. Evidence is *inadequate* to infer the presence or absence of a causal relationship (evidence that is sparse, of poor quality, or conflicting).
- D. Evidence is *suggestive of no causal relationship*.

Source: DHHS 2004.

evaluating evidence that an exposure sustained in military service had contributed to disease (IOM 2008); the report offers relevant coverage of the practice of causal inference.

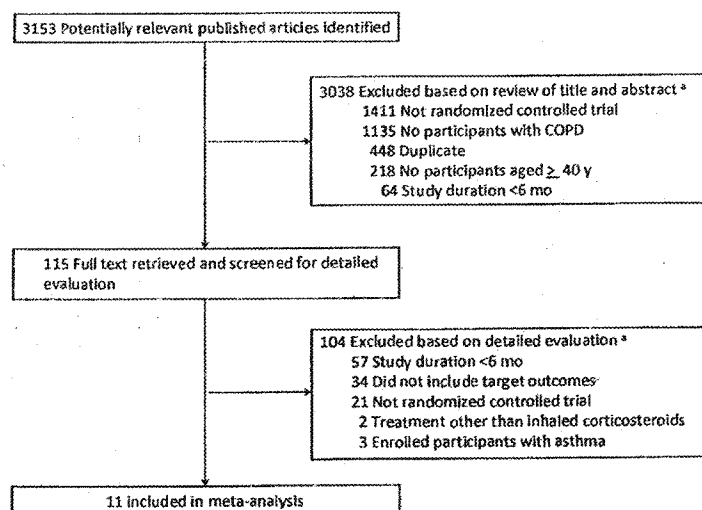
This brief and necessarily selective coverage of evidence reviews and evaluations shows that models are available that have proved successful in practice. They have several common elements: transparent and explicitly documented methods, consistent and critical evaluation of all relevant literature, application of a standardized approach for grading the strength of evidence, and clear and consistent summative language. Finally, highlighting features and limitations of the studies for use in quantitative assessments seems especially important for IRIS literature reviews.

A state-of-the-art literature review is essential for ensuring that the process of gathering evidence is comprehensive, transparent, and balanced. The committee suggests that EPA develop a detailed search strategy with search terms related to the specific questions that are addressed by the literature review. The yield of articles from searches can best be displayed graphically, documenting how initial search findings are narrowed to the articles in the final review selection on the basis of inclusion and exclusion criteria. Figure 7-3 provides an example of the selection process in a systematic review of a drug for lung disease. The progression from the initial 3,153 identified articles to the 11 reviewed is transparent. Although this example comes from an epidemiologic meta-analysis, a similar transparent process in which search terms, databases, and resources are listed and study selection is carefully tracked may be useful at all stages of the development of the IRIS assessment.

After studies are identified for review, the next step is to summarize the details and findings in evidence tables. Typically, such tables provide a link to the references, details of the study populations and methods, and key findings. They are prepared in a rigorous fashion with quality-assurance measures, such as using multiple abstractors (at least for a sample) and checking all numbers abstracted. If prepared correctly, the tables eliminate the need for long descriptions of studies and result in shorter text. Some draft IRIS assessments have begun to use a tabular format for systematic and concise presentation of evidence, and the committee encourages EPA to refine and expand that format as it revises the formaldehyde draft IRIS assessment and begins work on others.

The methods and findings of the studies are then evaluated with a standardized approach. Templates are useful for this purpose to ensure uniformity of approach, particularly if multiple reviewers are involved. Such standardized approaches are applied whether the research is epidemiologic (observational), experimental (randomized clinical trials), or toxicologic (animal bioassays). For example, for an observational epidemiologic study, a template for evaluation should consider the following:

- Approach used to identify the study population and the potential for selection bias.
- Study population characteristics and the generalizability of findings to other populations.
- Approach used for exposure assessment and the potential for information bias, whether differential (nonrandom) or nondifferential (random).
- Approach used for outcome identification and any potential bias.
- Appropriateness of analytic methods used.
- Potential for confounding to have influenced the findings.
- Precision of estimates of effect.
- Availability of an exposure metric that is used to model the severity of adverse response associated with a gradient of exposures.



**FIGURE 7-3** Example of an article-selection process. \*Articles could be excluded for more than one reason; therefore, summed exclusions exceed total. Abbreviation: COPD, chronic obstructive pulmonary disease. Source: Drummond et al. 2008. Reprinted with permission; copyright 2008, American Medical Association.

Similarly, a template for evaluation of a toxicology study in laboratory animals should consider the species and sex of animals studied, dosing information (dose spacing, dose duration, and route of exposure), end points considered, and the relevance of the end points to human end points of concern.

### Current Approaches to Hazard Identification

Hazard identification involves answering the question, Does the agent cause the adverse effect? (NRC 1983, 2009). Numerous approaches have been used for this purpose, and there is an extensive literature on causal inference, both on its philosophic underpinnings and on methods for evaluating the strength of evidence of causation. All approaches have in common a systematic identification of relevant evidence, criteria for evaluating the strength of evidence, and language for describing the strength of evidence of causation. The topic of causal inference and its role in decision-making was recently covered in the 2008 IOM report on evaluation of the presumptive decision-making process noted above. The 2004 report of the U.S. surgeon general on smoking and health (DHHS 2004) provided an updated review of the methods used in that series of reports.

The review approach for hazard identification embodies the elements described above and uses the criteria for evidence evaluation that have their origins in the 1964 report of the U.S. surgeon general (DHEW 1964) and the writings of Austin Bradford Hill, commonly known as the Hill criteria (see Table 7-1; Hill 1965). The criteria are not rigid and are not applied in a check-list manner; in fact, none is required for inferring a causal relationship, except for temporality inasmuch as exposure to the causal agent must precede the associated effect. The conclusion of causal inference is a clear statement on the strength of evidence of causation. For the purpose of hazard identification, such statements should follow a standardized classification to avoid ambiguity and to ensure comparability among different agents and outcomes.

Beyond the surgeon general's reports used here as an example, there are numerous examples of systematic approaches to hazard identification, including the monographs on carcinogenicity of the International Agency for Research on Cancer and the National Toxicology Program.<sup>1</sup> They have the same elements of systematic gathering and review of all lines of evidence and classification of the strength of evidence in a uniform and hierarchic structure.

#### Current Approaches to Dose-Response Assessment

The topic of dose-response assessment was covered in *Science and Decisions* (NRC 2009), which reviewed the current paradigm and called for a unified framework, bringing commonality to approaches for cancer and noncancer end points. That report also provides guidance on enhancing methods used to characterize uncertainty and variability. The present committee supports those recommendations but offers additional suggestions on the complementary coverage of the use of meta-analysis and pooled analysis in dose-response assessment.

IRIS assessments should address the following critical questions: Which studies should be included for derivation of reference values for noncancer outcomes and unit risks for cancer outcomes? Which dose-response models should be used for deriving those values? The latter question is related to model uncertainty in quantitative risk assessment and is not addressed here in this report. The former question is related to a fundamental issue of filtering the literature to identify the studies that provide the best dose-response information. A related question arises about how to combine information among studies because multiple studies may provide sufficient dose-response data. For this section, the committee assumes that the previously described evidence-based review has identified studies with adequate dose-response information to support some quantification of risk associated with exposure.

As suggested above, it would be unusual for a single study to trump all other studies providing information for setting reference values and unit risks. The combination of the analysis outcomes of different studies falls under the

---

<sup>1</sup>See <http://monographs.iarc.fr/index.php> and <http://ntp.niehs.nih.gov/>.

general description of meta-analysis (Normand 1999). The combination and synthesis of results of different studies appears central to an IRIS assessment, but such analyses require careful framing.

Stroup and colleagues (2000) provide a summary of recommendations for reporting meta-analyses of epidemiologic studies. Their proposal includes a table with a proposed check list that has broad categories for reporting, including background (such as problem definition and study population), search strategy (such as searchers, databases, and registries used), methods, results (such as graphic and tabular summaries, study description, and statistical uncertainty), discussion (such as bias and quality of included studies), and conclusion (such as generalization of conclusions and alternative explanations). Their recommendations on methods warrant specific consideration with reference to the development of an IRIS assessment, particularly those on evaluation and assessment of study relevance, rationale for selection and coding of studies, confounding, study quality, heterogeneity, and statistical methods. For the latter, key issues include the selection of models, the clarity with which findings are presented, and the availability of sufficient details to facilitate replication.

In combining study information, it is important that studies provide information on the same quantitative outcome, are conducted under similar conditions, and are of similar quality. If studies are of different quality, this might be addressed by weighting.

The simplest form of combining study information involves the aggregation of *p* values among a set of independent studies of the same null hypothesis. That simple approach might have appeal for establishing the relationship between some risk factor and an adverse outcome, but it is not useful for establishing exposure levels for a hazard. Thus, effect-size estimation among studies is usually of more interest for risk-estimation purposes and causality assessment. In this situation, a given effect is estimated for each study, and a combined estimate is obtained as a weighted average of study-specific effects in which the weights are inversely related to the precision associated with the estimation of each study-specific effect.

The question is whether EPA should routinely conduct meta-analysis for its IRIS assessments. Implicitly, the development of an IRIS assessment involves many of the steps associated with meta-analysis, including the collection and assessment of background literature. Assuming the availability of independent studies of the same end point and a comprehensive and unbiased inclusion of studies, questions addressed by a meta-analysis may be of great interest. Is there evidence of a homogeneous effect among studies? If not, can one understand the source of heterogeneity? If it is determined that a combined estimate is of interest (for example, an estimate of lifetime cancer risk based on combining study-specific estimates of this risk), a weighted estimate might be derived and reported.

**Case Study: Revision of the Approach to Evidence Review and Risk Assessment for National Ambient Air Quality Standards**

Approaches to evidence review and risk assessment vary within EPA. The recently revised approach used for NAAQSs offers an example that is particularly relevant because it represents a major change in an approach taken by one group in the National Center for Environmental Assessment. (EPA 2009b, 2010a,b)

Under Section 109 of the Clean Air Act, EPA is required to consider revisions of the NAAQSs for specified criteria air pollutants—currently particulate matter (PM), ozone, nitrogen dioxide, sulfur dioxide, carbon monoxide, and lead—every 5 years. Through 2009, the process for revision involved the development of two related documents that were both reviewed by the Clean Air Scientific Advisory Committee (CASAC) and made available for public comment. The first, the criteria document, was an encyclopedic compilation, sometimes several thousand pages long, of most scientific publications on the criteria pollutant that had been published since the previous review. Multiple authors contributed to the document, and there was generally little synthesis of the evidence, which was not accomplished in a systematic manner.

The other document was referred to as the staff paper. It was written by a different team in the Office of Air Quality Policy and Standards, and it identified the key scientific advances in the criteria document that were relevant to revising the NAAQSs. In the context of those advances, it offered the array of policy options around retaining or revising the NAAQSs that could be justified by recent research evidence. The linkages between the criteria document and the staff paper were general and not transparent.

The identified limitations of the process led to a proposal for its revision, and it took 2 years to complete the changes in the process. The new process replaces the criteria document with an integrated science assessment and a staff paper that includes a policy assessment. For the one pollutant, PM, that has nearly completed the full sequence, a risk and exposure analysis was also included.

The new documents address limitations of those used previously. The integrated science assessment is an evidence-based review that targets new studies as before. However, review methods are explicitly stated, and studies are reviewed in an informative and purposeful manner rather than in encyclopedic fashion. A main purpose of the integrated science assessment is to assess whether adverse health effects are causally linked to the pollutant under review. The integrated science assessment offers a five-category grading of strength of evidence on each outcome and follows the general weight-of-evidence approaches long used in public health. The intent is to base the risk and exposure analysis on effects for which causality is inferred or those at lower levels if they have particular public-health significance. The risk and exposure analysis brings

together the quantitative information on risk and exposure and provides estimates of the current burden of attributable morbidity and mortality and the estimates of avoidable and residual morbidity and mortality under various scenarios of changes in the NAAQS. Standard descriptors for uncertainty are now in place.

The policy assessment develops policy options on the basis of the findings of the integrated science assessment and the risk and exposure analysis. The policy assessment for the PM NAAQS is framed around a series of policy-relevant questions, such as, Does the available scientific evidence, as reflected in the integrated science assessment, support or call into question the adequacy of the protection afforded by the current 24-hr  $PM_{10}$  standard against effects associated with exposures to thoracic coarse particles? Evidence-based answers to the questions are provided with a reasonably standardized terminology for uncertainty.

For the most recent reassessment of the PM NAAQS, EPA staff and CASAC found the process to be effective; it led to greater transparency in evidence review and development of policy options than the prior process (Samet 2010). As noted above, the present committee sees the revision of the NAAQS review process as a useful example of how the agency was able to revise an entrenched process in a relatively short time.

### **Reframing the Development of the IRIS Assessment**

The committee was given the broad charge of reviewing the formaldehyde draft IRIS assessment and also asked to consider some specific questions. In addressing those questions, the committee found, as documented in Chapter 2, that some problems with the draft arose because of the processes and methods used to develop the assessment. Other committees have noted some of the same problems. Accordingly, the committee suggests here steps that EPA could take to improve IRIS assessment through the implementation of methods that would better reflect current practices. The committee offers a roadmap for changes in the development process if EPA concludes that such changes are needed. The term *roadmap* is used because the topics that need to be addressed are set out, but detailed guidance is not provided because that is seen as beyond the committee's charge. The committee's discussion of a reframing of the IRIS development process is based on its generic representation provided in Figure 7-2. The committee recognizes that the changes suggested would involve a multiyear process and extensive effort by the staff of the National Center for Environmental Assessment and input and review by the EPA Science Advisory Board and others. The recent revision of the NAAQS review process provides an example of an overhauling of an EPA evidence-review and risk-assessment process that took about 2 years.

In the judgment of the present and past committees, consideration needs to be given to how each step of the process could be improved and gains made in transparency and efficiency. Models for conducting IRIS reviews more effectively and efficiently are available. For each of the various components (Figure 7-2), methods have been developed, and there are exemplary approaches in assessments carried out elsewhere in EPA and by other organizations. In addition, there are relevant examples of evidence-based algorithms that EPA could draw on. Guidelines and protocols for the conduct of evidence-based reviews are available, as are guidelines for inference as to the strength of evidence of association and causation. Thus, EPA may be able to make changes in the assessment process relatively quickly by drawing on appropriate experts and selecting and adapting existing approaches.

One major, overarching issue is the use of weight of evidence in hazard identification. The committee recognizes that the terminology is embedded in various EPA guidelines (see Appendix B) and has proved useful. The determination of weight of evidence relies heavily on expert judgment. As called for by others, EPA might direct effort at better understanding how weight-of-evidence determinations are made with a goal of improving the process (White et al. 2009).

The committee highlights below what it considers critical for the development of a scientifically sound IRIS assessment. Although many elements are basic and have been addressed in the numerous EPA guidelines, implementation does not appear to be systematic or uniform in the development of the IRIS assessments.

#### **General Guidance for the Overall Process**

- Elaborate an overall, documented, and quality-controlled process for IRIS assessments.
- Ensure standardization of review and evaluation approaches among contributors and teams of contributors; for example, include standard approaches for reviews of various types of studies to ensure uniformity.
- Assess disciplinary structure of teams needed to conduct the assessments.

#### **Evidence Identification: Literature Collection and Collation Phase**

- Select outcomes on the basis of available evidence and understanding of mode of action.
- Establish standard protocols for evidence identification.
- Develop a template for description of the search approach.
- Use a database, such as the Health and Environmental Research Online (HERO) database, to capture study information and relevant quantitative data.

#### **Evidence Evaluation: Hazard Identification and Dose-Response Modeling**

- Standardize the presentation of reviewed studies in tabular or graphic form to capture the key dimensions of study characteristics, weight of evidence, and utility as a basis for deriving reference values and unit risks.
- Develop templates for evidence tables, forest plots, or other displays.
- Establish protocols for review of major types of studies, such as epidemiologic and bioassay.

#### **Weight-of-Evidence Evaluation: Synthesis of Evidence for Hazard Identification**

- Review use of existing weight-of-evidence guidelines.
- Standardize approach to using weight-of-evidence guidelines.
- Conduct agency workshops on approaches to implementing weight-of-evidence guidelines.
- Develop uniform language to describe strength of evidence on noncancer effects.
- Expand and harmonize the approach for characterizing uncertainty and variability.
- To the extent possible, unify consideration of outcomes around common modes of action rather than considering multiple outcomes separately.

#### **Selection of Studies for Derivation of Reference Values and Unit Risks**

- Establish clear guidelines for study selection.
  - Balance strengths and weaknesses.
  - Weigh human vs experimental evidence.
  - Determine whether combining estimates among studies is warranted.

#### **Calculation of Reference Values and Unit Risks**

- Describe and justify assumptions and models used. This step includes review of dosimetry models and the implications of the models for uncertainty factors; determination of appropriate points of departure (such as benchmark dose, no-observed-adverse-effect level, and lowest observed-adverse-effect level), and assessment of the analyses that underlie the points of departure.
- Provide explanation of the risk-estimation modeling processes (for example, a statistical or biologic model fit to the data) that are used to develop a unit risk estimate.

- Assess the sensitivity of derived estimates to model assumptions and end points selected. This step should include appropriate tabular and graphic displays to illustrate the range of the estimates and the effect of uncertainty factors on the estimates.
- Provide adequate documentation for conclusions and estimation of reference values and unit risks. As noted by the committee throughout the present report, sufficient support for conclusions in the formaldehyde draft IRIS assessment is often lacking. Given that the development of specific IRIS assessments and their conclusions are of interest to many stakeholders, it is important that they provide sufficient references and supporting documentation for their conclusions. Detailed appendixes, which might be made available only electronically, should be provided when appropriate.

## REFERENCES

- DHEW (U.S. Department of Health Education and Welfare). 1964. Smoking and Health. Report of the Advisory Committee to the Surgeon General. Public Health Service Publication No. 1103. Washington, DC: U.S. Government Printing Office [online]. Available: [http://profiles.nlm.nih.gov/NN/B/B/M/Q/\\_/nnbbmq.pdf](http://profiles.nlm.nih.gov/NN/B/B/M/Q/_/nnbbmq.pdf) [accessed Feb. 1, 2011].
- DHHS (U.S. Department of Health and Human Services). 2004. The Health Consequences of Smoking: A Report of the Surgeon General. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA [online]. Available: [http://www.cdc.gov/tobacco/data\\_statistics/sgr/2004/complete\\_report/index.htm](http://www.cdc.gov/tobacco/data_statistics/sgr/2004/complete_report/index.htm) [accessed Nov. 22, 2010].
- Drummond, M.B., E.C. Dasenbrook, M.W. Pitz, D.J. Murphy, and E. Fan. 2008. Inhaled corticosteroids in patients with stable chronic obstructive pulmonary disease: A systematic review and meta-analysis. *JAMA*. 300(20):2407-2416.
- EPA (U.S. Environmental Protection Agency). 2009a. New Process for Development of Integrated Risk Information System Health Assessments. Memorandum to Assistant Administrators, General Counsel, Inspector General, Chief Financial Officer, Chief of Staff, Associate Administrators, and Regional Administrators, from Lisa P. Jackson, the Administrator, U.S. Environmental Protection Agency, Washington, DC. May 21, 2009 [online]. Available: [http://www.epa.gov/iris/pdfs/IRIS\\_PROCESS\\_MEMO.5.21.09.PDF](http://www.epa.gov/iris/pdfs/IRIS_PROCESS_MEMO.5.21.09.PDF) [accessed Nov. 23, 2010].
- EPA (U.S. Environmental Protection Agency). 2009b. Integrated Science Assessment for Particulate Matter (Final Report). EPA/600/R-08/139F. National Center for Environmental Assessment-RTP Division, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC. December 2009 [online]. Available: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546> [accessed March 2, 2011].
- EPA (U.S. Environmental Protection Agency). 2010a. Quantitative Health Risk Assessment for Particulate Matter (Final Report). EPA-452/R-10-005. Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. June 2010

## Appendix B

### Weight-of-Evidence Descriptions from U.S. Environmental Protection Agency Guidelines

The text in this appendix was excerpted directly from the indicated guidelines of the U.S. Environmental Protection Agency (EPA).

#### GUIDELINES FOR MUTAGENICITY RISK ASSESSMENT

The evidence for a chemical's ability to produce mutations and to interact with the germinal target is integrated into a weight-of-evidence judgment that the agent may pose a hazard as a potential human germ-cell mutagen. All information bearing on the subject, whether indicative of potential concern or not, must be evaluated. Whatever evidence may exist from humans must also be factored into the assessment.

All germ-cell stages are important in evaluating chemicals because some chemicals have been shown to be positive in postgonial stages but not in gonial (Russell et al., 1984). When human exposures occur, effects on postgonial stages should be weighted by the relative sensitivity and the duration of the stages. Chemicals may show positive effects for some endpoints and in some test systems, but negative responses in others. Each review must take into account the limitations in the testing and in the types of responses that may exist.

To provide guidance as to the categorization of the weight of evidence, a classification scheme is presented to illustrate, in a simplified sense, the strength of the information bearing on the potential for human germ-cell mutagenicity. It is not possible to illustrate all potential combinations of evidence, and considerable judgment must be exercised in reaching conclusions. In addition, certain responses in tests that do not measure direct mutagenic end points (e.g., SCE induction in mammalian germ cells) may provide a basis for raising the weight

of evidence from one category to another. The categories are presented in decreasing order of strength of evidence.

1. Positive data derived from human germ-cell mutagenicity studies, when available, will constitute the highest level of evidence for human mutagenicity.
2. Valid positive results from studies on heritable mutational events (of any kind) in mammalian germ cells.
3. Valid positive results from mammalian germ-cell chromosome aberration studies that do not include an intergeneration test.
4. Sufficient evidence for a chemical's interaction with mammalian germ cells, together with valid positive mutagenicity test results from two assay systems, at least one of which is mammalian (in vitro or in vivo). The positive results may both be for gene mutations or both for chromosome aberrations; if one is for gene mutations and the other for chromosome aberrations, both must be from mammalian systems.
5. Suggestive evidence for a chemical's interaction with mammalian germ cells, together with valid positive mutagenicity evidence from two assay systems as described under 4, above. Alternatively, positive mutagenicity evidence of less strength than defined under 4, above, when combined with sufficient evidence for a chemical's interaction with mammalian germ cells.
6. Positive mutagenicity test results of less strength than defined under 4, combined with suggestive evidence for a chemical's interaction with mammalian germ cells.
7. Although definitive proof of nonmutagenicity is not possible, a chemical could be classified operationally as a nonmutagen for human germ cells if it gives valid negative test results for all endpoints of concern.
8. Inadequate evidence bearing on either mutagenicity or chemical interaction with mammalian germ cells (EPA 1986, Pp 9-10).

#### **METHODS FOR DERIVATION OF INHALATION REFERENCE CONCENTRATIONS AND APPLICATION OF INHALATION DOSIMETRY**

The culmination of the hazard identification phase of any risk assessment involves integrating a diverse data collection into a cohesive, biologically plausible toxicity "picture"; that is, to develop the weight of evidence that the chemical poses a hazard to humans. The salient points from each of the laboratory animal and human studies in the entire data base should be summarized as should the analysis devoted to examining the variation or consistency among factors (usually related to the mechanism of action), in order to establish the likely outcome for exposure to this chemical. From this analysis, an appropriate animal model or additional factors pertinent to human extrapolation may be identified.

The utility of a given study is often related to the nature and quality of the other available data. For example, clinical pharmacokinetic studies may validate that the target organ or disease in laboratory animals is likely to be the same effect observed in the exposed human population. However, if a cohort study describing the nature of the dose-response relationship were available, the clinical description would rarely give additional information. An apparent conflict may arise in the analysis when an association is observed in toxicologic but not epidemiologic data, or vice versa. The analysis then should focus on reasons for the apparent difference in order to resolve the discrepancy. For example, the epidemiologic data may have contained other exposures not accounted for, or the laboratory animal species tested may have been inappropriate for the mechanism of action. A framework for approaching data summary is provided in Table 2-6. Table 2-7 provides the specific uses of various types of human data in such an approach. These guidelines have evolved from criteria used to establish causal significance, such as those developed by the American Thoracic Society (1985) to assess the causal significance of an air toxicant and a health effect. The criteria for establishing causal significance can be found in Appendix C. In general, the following factors enhance the weight of evidence on a chemical:

- Clear evidence of a dose-response relationship;
- Similar effects across sex, strain, species, exposure routes, or in multiple experiments;
- Biologically plausible relationship between metabolism data, the postulated mechanism of action, and the effect of concern;
- Similar toxicity exhibited by structurally related compounds;
- Some correlation between the observed chemical toxicity and human evidence.

The greater the weight of evidence, the greater the confidence in the conclusion derived. Developing improved weight-of-evidence schemes for various noncancer health effect categories has been the focus of efforts by the Agency to improve health risk assessment methodologies (Perlin and McCormack, 1988).

Another difficulty encountered in this summarizing process is that certain studies may produce apparently positive or negative results, yet may be flawed. The flaws may have arisen from inappropriate design or execution in performance (e.g., lack of statistical power or adjustment of dosage during the course of the study to avoid undesirable toxic effects). The treatment of flawed results is critical; although there is something to be learned from every study, the extent that a study should be used is dependent on the nature of the flaw (Society of Toxicology, 1982). A flawed negative study could only provide a false sense of security, whereas a flawed positive study may contribute to some limited understanding. Although there is no substitute for good science, grey areas such as this are ultimately a matter of scientific judgment. The risk assessor will have to decide what is and is not useful within the framework outlined earlier.

Studies meeting the criteria detailed in Sections 2.1.1 and 2.1.2 (epidemiologic, nonepidemiologic data), and experimental studies on laboratory animals that fit into this weight-of-evidence framework are used in the quantitative dose-response assessment discussed in Chapter 4 (EPA 1994, Pp 2-42 to 2-46).

#### GUIDELINES FOR DEVELOPMENTAL TOXICITY RISK ASSESSMENT

The 1989 Proposed Amendments described important considerations in determining the relative weight of various kinds of data in estimating the risk of developmental toxicity in humans. The intent of the proposed weight-of-evidence (WOE) scheme was that it not be used in isolation, but be used as the first step in the risk assessment process, to be integrated with dose-response information and the exposure assessment.

The WOE scheme was the subject of a considerable number of public comments, and was one of the major concerns of the SAB. The concern of public commentators was that the reference to human developmental toxicity in this scheme suggested that a chemical could be prematurely designated, and perhaps labeled, as causing developmental toxicity in humans prior to the completion of the risk assessment process. The SAB suggested that the intended use of this scheme was not consistent with the use of the term "weight of evidence" in other contexts, since WOE is usually thought of as an evaluation of the total composite of information available to make a judgment about risk. In addition, the SAB Committee proposed that the Agency consider development of a more conceptual approach using decision analytical techniques to predict the relationships among various outcomes.

In the final Guidelines, the terminology used in the WOE scheme has been completely changed and retitled "Characterization of the Health-Related Database." The intended purpose of the scheme is to provide a framework and criteria for making a decision on whether or not sufficient data are available to conduct a risk assessment. This decision is based on the available data, whether animal or human, and does not necessarily imply human hazard. This decision process is part of, but not the complete, WOE evaluation, which also takes into account the RfDDT or RfCDT and the human exposure information, culminating in risk characterization.

The final Guidelines also place strong emphasis on the integration of the dose-response evaluation with hazard information in characterizing the sufficiency of the health-related database. In line with this approach, the Guidelines have been reorganized to combine hazard identification and dose-response evaluation. Finally, the SAB comments on developing a conceptual matrix provide an interesting challenge, but current data indicate that the relationships among endpoints of developmental toxicity are not consistent across chemicals or species. The Agency is currently supporting modeling efforts to further explore the relationship among various development toxicity endpoints and the

development of biologically based dose-response models that consider multiple effects (EPA 1991, Pp 69-70).

#### A REVIEW OF THE REFERENCE DOSE AND REFERENCE CONCENTRATION PROCESSES

A weight-of-evidence approach such as that provided in EPA's RfC Methodology (U.S. EPA, 1994) or in EPA's proposed guidelines for carcinogen risk assessment (U.S. EPA, 1999a) should be used in assessing the database for an agent. This approach requires a critical evaluation of the entire body of available data for consistency and biological plausibility. Potentially relevant studies should be judged for quality and studies of high quality given much more weight than those of lower quality. When both epidemiological and experimental data are available, similarity of effects between humans and animals is given more weight. If the mechanism or mode of action is well characterized, this information is used in the interpretation of observed effects in either human or animal studies. Weight of evidence is not to be interpreted as simply tallying the number of positive and negative studies, nor does it imply an averaging of the doses or exposures identified in individual studies that may be suitable as points of departure (PODs) for risk assessment. The study or studies used for the POD are identified by an informed and expert evaluation of all the available evidence (EPA 2002b, Pp 4-11 to 4-12).

#### GUIDELINES FOR CARCINOGEN RISK ASSESSMENT

The cancer guidelines emphasize the importance of weighing all of the evidence in reaching conclusions about the human carcinogenic potential of agents. This is accomplished in a single integrative step after assessing all of the individual lines of evidence, which is in contrast to the step-wise approach in the 1986 cancer guidelines. Evidence considered includes tumor findings, or lack thereof, in humans and laboratory animals; an agent's chemical and physical properties; its structure-activity relationships (SARs) as compared with other carcinogenic agents; and studies addressing potential carcinogenic processes and mode(s) of action, either *in vivo* or *in vitro*. Data from epidemiologic studies are generally preferred for characterizing human cancer hazard and risk. However, all of the information discussed above could provide valuable insights into the possible mode(s) of action and likelihood of human cancer hazard and risk. The cancer guidelines recognize the growing sophistication of research methods, particularly in their ability to reveal the modes of action of carcinogenic agents at cellular and subcellular levels as well as toxicokinetic processes.

Weighing of the evidence includes addressing not only the likelihood of human carcinogenic effects of the agent but also the conditions under which such effects may be expressed, to the extent that these are revealed in the toxicological and other biologically important features of the agent.

The weight of evidence narrative to characterize hazard summarizes the results of the hazard assessment and provides a conclusion with regard to human carcinogenic potential. The narrative explains the kinds of evidence available and how they fit together in drawing conclusions, and it points out significant issues/strengths/limitations of the data and conclusions. Because the narrative also summarizes the mode of action information, it sets the stage for the discussion of the rationale underlying a recommended approach to dose-response assessment.

In order to provide some measure of clarity and consistency in an otherwise free-form, narrative characterization, standard descriptors are used as part of the hazard narrative to express the conclusion regarding the weight of evidence for carcinogenic hazard potential. There are five recommended standard hazard descriptors: "*Carcinogenic to Humans*," "*Likely to Be Carcinogenic to Humans*," "*Suggestive Evidence of Carcinogenic Potential*," "*Inadequate Information to Assess Carcinogenic Potential*," and "*Not Likely to Be Carcinogenic to Humans*." Each standard descriptor may be applicable to a wide variety of data sets and weights of evidence and is presented only in the context of a weight of evidence narrative. Furthermore, as described in Section 2.5 of these cancer guidelines, more than one conclusion may be reached for an agent (EPA 2005b, Pp 1-11 to 1-12).

The *weight of evidence narrative* is a short summary (one to two pages) that explains an agent's human carcinogenic potential and the conditions that characterize its expression. It should be sufficiently complete to be able to stand alone, highlighting the key issues and decisions that were the basis for the evaluation of the agent's potential hazard. It should be sufficiently clear and transparent to be useful to risk managers and non-expert readers. It may be useful to summarize all of the significant components and conclusions in the first paragraph of the narrative and to explain complex issues in more depth in the rest of the narrative.

The weight of the evidence should be presented as a narrative laying out the complexity of information that is essential to understanding the hazard and its dependence on the quality, quantity, and type(s) of data available, as well as the circumstances of exposure or the traits of an exposed population that may be required for expression of cancer. For example, the narrative can clearly state to what extent the determination was based on data from human exposure, from animal experiments, from some combination of the two, or from other data. Similarly, information on mode of action can specify to what extent the data are from *in vivo* or *in vitro* exposures or based on similarities to other chemicals. The extent to which an agent's mode of action occurs only on reaching a minimum dose or a minimum duration should also be presented. A hazard might also be expressed disproportionately in individuals possessing a specific gene; such characterizations may follow from a better understanding of the human genome. Furthermore, route of exposure should be used to qualify a hazard if, for example, an agent is not absorbed by some routes. Similarly, a hazard can be attribut-

able to exposures during a susceptible lifestage on the basis of our understanding of human development.

The weight of evidence-of-evidence narrative should highlight:

- the quality and quantity of the data;
- all key decisions and the basis for these major decisions; and
- any data, analyses, or assumptions that are unusual for or new to EPA.

To capture this complexity, a weight of evidence narrative generally includes

- conclusions about human carcinogenic potential (choice of descriptor(s), described below),
- a summary of the key evidence supporting these conclusions (for each descriptor used), including information on the type(s) of data (human and/or animal, *in vivo* and/or *in vitro*) used to support the conclusion(s),
- available information on the epidemiologic or experimental conditions that characterize expression of carcinogenicity (e.g., if carcinogenicity is possible only by one exposure route or only above a certain human exposure level),
- a summary of potential modes of action and how they reinforce the conclusions,
- indications of any susceptible populations or lifestages, when available, and
- a summary of the key default options invoked when the available information is inconclusive.

To provide some measure of clarity and consistency in an otherwise free-form narrative, the weight of evidence descriptors are included in the first sentence of the narrative. Choosing a descriptor is a matter of judgment and cannot be reduced to a formula. Each descriptor may be applicable to a wide variety of potential data sets and weights of evidence. These descriptors and narratives are intended to permit sufficient flexibility to accommodate new scientific understanding and new testing methods as they are developed and accepted by the scientific community and the public. Descriptors represent points along a continuum of evidence; consequently, there are gradations and borderline cases that are clarified by the full narrative. Descriptors, as well as an introductory paragraph, are a short summary of the complete narrative that preserves the complexity that is an essential part of the hazard characterization. Users of these cancer guidelines and of the risk assessments that result from the use of these cancer guidelines should consider the entire range of information included in the narrative rather than focusing simply on the descriptor.

In borderline cases, the narrative explains the case for choosing one descriptor and discusses the arguments for considering but not choosing another.

For example, between "suggestive" and "likely" or between "suggestive" and "inadequate," the explanation clearly communicates the information needed to consider appropriately the agent's carcinogenic potential in subsequent decisions.

Multiple descriptors can be used for a single agent, for example, when carcinogenesis is dose- or route-dependent. For example, if an agent causes point-of-contact tumors by one exposure route but adequate testing is negative by another route, then the agent could be described as likely to be carcinogenic by the first route but not likely to be carcinogenic by the second. Another example is when the mode of action is sufficiently understood to conclude that a key event in tumor development would not occur below a certain dose range. In this case, the agent could be described as likely to be carcinogenic above a certain dose range but not likely to be carcinogenic below that range.

Descriptors can be selected for an agent that has not been tested in a cancer bioassay if sufficient other information, e.g., toxicokinetic and mode of action information, is available to make a strong, convincing, and logical case through scientific inference. For example, if an agent is one of a well-defined class of agents that are understood to operate through a common mode of action and if that agent has the same mode of action, then in the narrative the untested agent would have the same descriptor as the class. Another example is when an untested agent's effects are understood to be caused by a human metabolite, in which case in the narrative the untested agent could have the same descriptor as the metabolite. As new testing methods are developed and used, assessments may increasingly be based on inferences from toxicokinetic and mode of action information in the absence of tumor studies in animals or humans.

When a well-studied agent produces tumors only at a point of initial contact, the descriptor generally applies only to the exposure route producing tumors unless the mode of action is relevant to other routes. The rationale for this conclusion would be explained in the narrative.

When tumors occur at a site other than the point of initial contact, the descriptor generally applies to all exposure routes that have not been adequately tested at sufficient doses. An exception occurs when there is convincing information, e.g., toxicokinetic data that absorption does not occur by another route.

When the response differs qualitatively as well as quantitatively with dose, this information should be part of the characterization of the hazard. In some cases reaching a certain dose range can be a precondition for effects to occur, as when cancer is secondary to another toxic effect that appears only above a certain dose. In other cases exposure duration can be a precondition for hazard if effects occur only after exposure is sustained for a certain duration. These considerations differ from the issues of relative absorption or potency at different dose levels because they may represent a discontinuity in a dose-response function.

When multiple bioassays are inconclusive, mode of action data are likely to hold the key to resolution of the more appropriate descriptor. When bioassays

are few, further bioassays to replicate a study's results or to investigate the potential for effects in another sex, strain, or species may be useful.

When there are few pertinent data, the descriptor makes a statement about the database, for example, "Inadequate Information to Assess Carcinogenic Potential," or a database that provides "Suggestive Evidence of Carcinogenic Potential." With more information, the descriptor expresses a conclusion about the agent's carcinogenic potential to humans. If the conclusion is positive, the agent could be described as "Likely to Be Carcinogenic to Humans" or, with strong evidence, "Carcinogenic to Humans." If the conclusion is negative, the agent could be described as "Not Likely to Be Carcinogenic to Humans."

Although the term "likely" can have a probabilistic connotation in other contexts, its use as a weight of evidence descriptor does not correspond to a quantifiable probability of whether the chemical is carcinogenic. This is because the data that support cancer assessments generally are not suitable for numerical calculations of the probability that an agent is a carcinogen. Other health agencies have expressed a comparable weight of evidence using terms such as "Reasonably Anticipated to Be a Human Carcinogen" (NTP) or "Probably Carcinogenic to Humans" (International Agency for Research on Cancer).

The following descriptors can be used as an introduction to the weight of evidence narrative. The examples presented in the discussion of the descriptors are illustrative. The examples are neither a checklist nor a limitation for the descriptor. The complete weight of evidence narrative, rather than the descriptor alone, provides the conclusions and the basis for them.

#### ***"Carcinogenic to Humans"***

This descriptor indicates strong evidence of human carcinogenicity. It covers different combinations of evidence.

- This descriptor is appropriate when there is convincing epidemiologic evidence of a causal association between human exposure and cancer.
- Exceptionally, this descriptor may be equally appropriate with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence. It can be used when all of the following conditions are met: (a) there is strong evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action but not enough for a causal association, and (b) there is extensive evidence of carcinogenicity in animals, and (c) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals, and (d) there is strong evidence that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information. In this case, the narrative includes a summary of both the experimental and epidemiologic information on mode of action and also an indication of the

relative weight that each source of information carries, e.g., based on human information, based on limited human and extensive animal experiments.

*"Likely to Be Carcinogenic to Humans"*

This descriptor is appropriate when the weight of the evidence is adequate to demonstrate carcinogenic potential to humans but does not reach the weight of evidence for the descriptor "Carcinogenic to Humans." Adequate evidence consistent with this descriptor covers a broad spectrum. As stated previously, the use of the term "likely" as a weight of evidence descriptor does not correspond to a quantifiable probability. The examples below are meant to represent the broad range of data combinations that are covered by this descriptor; they are illustrative and provide neither a checklist nor a limitation for the data that might support use of this descriptor. Moreover, additional information, e.g., on mode of action, might change the choice of descriptor for the illustrated examples. Supporting data for this descriptor may include:

- an agent demonstrating a plausible (but not definitively causal) association between human exposure and cancer, in most cases with some supporting biological, experimental evidence, though not necessarily carcinogenicity data from animal experiments;
- an agent that has tested positive in animal experiments in more than one species, sex, strain, site, or exposure route, with or without evidence of carcinogenicity in humans;
- a positive tumor study that raises additional biological concerns beyond that of a statistically significant result, for example, a high degree of malignancy, or an early age at onset;
- a rare animal tumor response in a single experiment that is assumed to be relevant to humans; or
- a positive tumor study that is strengthened by other lines of evidence, for example, either plausible (but not definitively causal) association between human exposure and cancer or evidence that the agent or an important metabolite causes events generally known to be associated with tumor formation (such as DNA reactivity or effects on cell growth control) likely to be related to the tumor response in this case.

*"Suggestive Evidence of Carcinogenic Potential"*

This descriptor of the database is appropriate when the weight of evidence is suggestive of carcinogenicity; a concern for potential carcinogenic effects in humans is raised, but the data are judged not sufficient for a stronger conclusion. This descriptor covers a spectrum of evidence associated with varying levels of concern for carcinogenicity, ranging from a positive cancer result in the only

study on an agent to a single positive cancer result in an extensive database that includes negative studies in other species. Depending on the extent of the database, additional studies may or may not provide further insights. Some examples include:

- a small, and possibly not statistically significant, increase in tumor incidence observed in a single animal or human study that does not reach the weight of evidence for the descriptor "Likely to Be Carcinogenic to Humans." The study generally would not be contradicted by other studies of equal quality in the same population group or experimental system (see discussions of *conflicting evidence* and *differing results*, below);
- a small increase in a tumor with a high background rate in that sex and strain, when there is some but insufficient evidence that the observed tumors may be due to intrinsic factors that cause background tumors and not due to the agent being assessed. (When there is a high background rate of a specific tumor in animals of a particular sex and strain, then there may be biological factors operating independently of the agent being assessed that could be responsible for the development of the observed tumors.) In this case, the reasons for determining that the tumors are not due to the agent are explained;
- evidence of a positive response in a study whose power, design, or conduct limits the ability to draw a confident conclusion (but does not make the study fatally flawed), but where the carcinogenic potential is strengthened by other lines of evidence (such as structure-activity relationships); or
- a statistically significant increase at one dose only, but no significant response at the other doses and no overall trend.

#### *"Inadequate Information to Assess Carcinogenic Potential"*

This descriptor of the database is appropriate when available data are judged inadequate for applying one of the other descriptors. Additional studies generally would be expected to provide further insights. Some examples include:

- little or no pertinent information;
- conflicting evidence, that is, some studies provide evidence of carcinogenicity but other studies of equal quality in the same sex and strain are negative. *Differing results*, that is, positive results in some studies and negative results in one or more different experimental systems, do not constitute *conflicting evidence*, as the term is used here. Depending on the overall weight of evidence, differing results can be considered either suggestive evidence or likely evidence; or negative results that are not sufficiently robust for the descriptor, "Not Likely to Be Carcinogenic to Humans."

- negative results that are not sufficiently robust for the descriptor, "Not Likely to Be Carcinogenic to Humans."

*"Not Likely to Be Carcinogenic to Humans"*

This descriptor is appropriate when the available data are considered robust for deciding that there is no basis for human hazard concern. In some instances, there can be positive results in experimental animals when there is strong, consistent evidence that each mode of action in experimental animals does not operate in humans. In other cases, there can be convincing evidence in both humans and animals that the agent is not carcinogenic. The judgment may be based on data such as:

- animal evidence that demonstrates lack of carcinogenic effect in both sexes in well designed and well-conducted studies in at least two appropriate animal species (in the absence of other animal or human data suggesting a potential for cancer effects),
- convincing and extensive experimental evidence showing that the only carcinogenic effects observed in animals are not relevant to humans,
- convincing evidence that carcinogenic effects are not likely by a particular exposure route (see Section 2.3), or
- convincing evidence that carcinogenic effects are not likely below a defined dose range. A descriptor of "not likely" applies only to the circumstances supported by the data. For example, an agent may be "Not Likely to Be Carcinogenic" by one route but not necessarily by another. In those cases that have positive animal experiment(s) but the results are judged to be not relevant to humans, the narrative discusses why the results are not relevant.

*Multiple Descriptors*

More than one descriptor can be used when an agent's effects differ by dose or exposure route. For example, an agent may be "Carcinogenic to Humans" by one exposure route but "Not Likely to Be Carcinogenic" by a route by which it is not absorbed. Also, an agent could be "Likely to Be Carcinogenic" above a specified dose but "Not Likely to Be Carcinogenic" below that dose because a key event in tumor formation does not occur below that dose (EPA 2005b, Pp 2-49 to 2-58).

**A FRAMEWORK FOR ASSESSING HEALTH RISKS OF  
ENVIRONMENTAL EXPOSURES TO CHILDREN**

The WOE approach requires a critical evaluation (expert judgment) of all available data for consistency and biological plausibility. Criteria for this as-

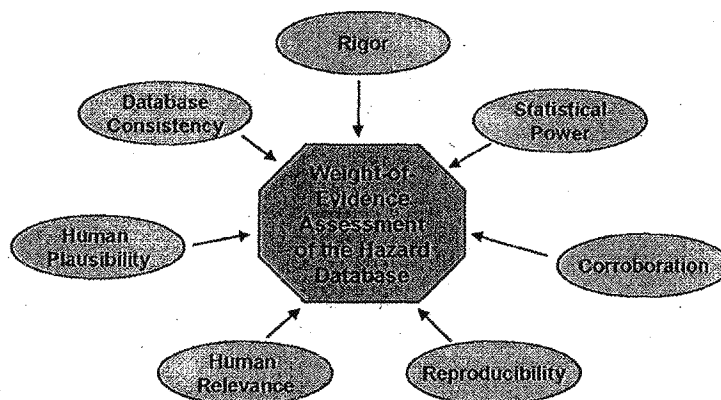
assessment are not presented here; rather, considerations important for the WOE are described. The key to WOE conclusions is the provision of a clear justification for decisions. Finally, the extent of the database is summarized, and assumptions made in the assessment are explicitly detailed. Further details about EPA's WOE approach can be found in the *Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry* (U.S. EPA, 1994), *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005b), and *Supplemental Guidance for Assessing Cancer Susceptibility from Early Life Exposure to Carcinogens* (U.S. EPA, 2005c). *A Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002b, Section 4.3.2.1.) and *Determination of the Appropriate FQPA Safety Factor(s) on Tolerance Assessment* (U.S. EPA, 2002c, Section III) provide additional detail on the WOE.

Key themes for the consideration of toxicity data in a WOE assessment, as adapted from Gray et al. (2001), are shown in Figure 4-5. This figure focuses on judging animal studies within a WOE assessment. However, if adequate human studies are available they would be given more weight. The process for evaluating these considerations is described in the following subsections. In this process, the quality of potentially relevant studies is judged, modifiers and interactions are detailed, outcomes across species are compared, TK and TD data are examined and weighed for comparisons across species, and the uncertainties and data gaps are determined. SARs with other chemicals or chemical classes are explored to determine the extent to which these data can inform the assessment via an MOA discussion or reduce uncertainties.

#### GUIDELINES FOR NEUROTOXICITY RISK ASSESSMENT

The interpretation of data as indicative of a potential neurotoxic effect involves the evaluation of the validity of the database. This approach and these terms have been adapted from the literature on human psychological testing (Sette, 1987; Sette and MacPhail, 1992), where they have long been used to evaluate the level of confidence in different measures of intelligence or other abilities, aptitudes, or feelings. There are four principal questions that should be addressed: whether the effects result from exposure (content validity); whether the effects are adverse or toxicologically significant (construct validity); whether there are correlative measures among behavioral, physiological, neurochemical, and morphological endpoints (concurrent validity); and whether the effects are predictive of what will happen under various conditions (predictive validity). Addressing these issues can provide a useful framework for evaluating either human or animal studies or the weight of evidence for a chemical (Sette, 1987; Sette and MacPhail, 1992). The next sections indicate the extent to which chemically induced changes can be interpreted as providing evidence of neurotoxicity.

The qualitative characterization of neurotoxic hazard can be based on either human or animal data (Anger, 1984; Reiter, 1987; U.S. EPA, 1994). Such data can result from accidental, inappropriate, or controlled experimental exposures. This section describes many of the general and some of the specific characteristics of human studies and reports of neurotoxicity. It then describes some features of animal studies of neuroanatomical, neurochemical, neurophysiological, and behavioral effects relevant to risk assessment. The process of characterizing the sufficiency or insufficiency of neurotoxic effects for risk assessment is described in section 3.3. Additional sources of information relevant to hazard characterization, such as comparisons of molecular structure among compounds and in vitro screening methods, are also discussed.



**FIGURE 4-5** Conceptual view of a weight of evidence (WOE) assessment. This figure illustrates the critical considerations within a WOE assessment of toxicity data. *Rigor* is the degree of proper conduct and analysis of a study; greater weight is generally given to more rigorous studies. *Statistical Power* is the ability of a study to detect effects of a given magnitude. *Corroboration* means that specific effects are replicated in similar studies, similar effects are observed under varied conditions and/or similar effects are observed in multiple laboratories. *Reproducibility* means that an effect is observed in multiple species by various routes of exposure. *Relevance to Humans* means that similar effects are observed in humans or in a species taxonomically related to humans or at doses similar to those expected in humans. *Plausibility to Humans* is the determination of whether a similar metabolism, mechanisms of damage and repair, and molecular target of response could be expected to occur in humans, based on an evaluation of the biologic mechanism of a toxic response in animals. *Database Consistency* is the extent to which all of the data are similar in outcome and dose (exposure-response) and are operating under a single biologically plausible assumption (mode of action). Source: Adapted from Gray et al. 2001, EPA 2006, Pp 29-30.

The hazard characterization should:

- a. Identify strengths and limitations of the database:
  - Epidemiological studies (case reports, cross-sectional, case-control, cohort, or human laboratory exposure studies);
  - Animal studies (including structural or neuropathological, neurochemical, neurophysiological, behavioral or neurological, or developmental endpoints).
- b. Evaluate the validity of the database:
  - Content validity (effects result from exposure);
  - Construct validity (effects are adverse or toxicologically significant);
  - Concurrent validity (correlative measures among behavioral, physiological, neurochemical, or morphological endpoints);
  - Predictive validity (effects are predictive of what will happen under various conditions).
- c. Identify and describe key toxicological studies.
- d. Describe the type of effects:
  - Structural (neuroanatomical alternations);
  - Functional (neurochemical, neurophysiological, behavioral alterations).
- e. Describe the nature of the effects (irreversible, reversible, transient, progressive, delayed, residual, or latent).
- f. Describe how much is known about how (through what biological mechanism) the chemical produces adverse effects.
- g. Discuss other health endpoints of concern.
- h. Comment on any nonpositive data in humans or animals.
- i. Discuss the dose-response data (epidemiological or animal) available for further dose-response analysis.
- j. Discuss the route, level, timing, and duration of exposure in studies demonstrating neurotoxicity as compared to expected human exposures.
- k. Summarize the hazard characterization:
  - Confidence in conclusions;
  - Alternative conclusions also supported by the data;
  - Significant data gaps; and
  - Highlights of major assumptions.

## REFERENCES

- American Thoracic Society. 1985. Guidelines as to what constitutes an adverse respiratory health effect, with special reference to epidemiologic studies of air pollution. *Am. Rev. Respir. Dis.* 131(4):666-668.

- Anger, W.K. 1984. Neurobehavioral testing of chemicals: Impact on recommended standards. *Neurobehav. Toxicol. Teratol.* 6(2):147-153.
- EPA (U.S. Environmental Protection Agency). 1986. Guidelines for Mutagenicity Risk Assessment. U.S. Environmental Protection Agency [online]. Available: <http://www.epa.gov/osa/mmoaframework/pdfs/MUTAGEN2.PDF> [accessed Nov. 19, 2010].
- EPA (U.S. Environmental Protection Agency). 1991. Guidelines for Developmental Toxicity Risk Assessment. U.S. Environmental Protection Agency [online]. Available: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=23162> [accessed Nov. 19, 2010].
- EPA (U.S. Environmental Protection Agency). 1994. Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. U.S. Environmental Protection Agency [online]. Available: <http://www.epa.gov/raf/publications/methods-derivation-inhalation-ref.htm> [accessed Nov. 19, 2010].
- EPA (U.S. Environmental Protection Agency). 1998. Guidelines for Neurotoxicity Risk Assessment. U.S. Environmental Protection Agency [online]. Available: <http://www.epa.gov/raf/publications/pdfs/NEUROTOX.PDF> [accessed Dec. 16, 2010].
- EPA (U.S. Environmental Protection Agency). 1999a. Guidelines for Carcinogen Risk Assessment [Review Draft]. NCEA-F-0644. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. July 1999 [online]. Available: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=54932#Download> [accessed Mar. 17, 2011].
- EPA (U.S. Environmental Protection Agency). 2002b. A Review of the Reference Dose and Reference Concentration Process. U.S. Environmental Protection Agency [online]. Available: [http://www.epa.gov/iris/RFD\\_FINAL1.pdf](http://www.epa.gov/iris/RFD_FINAL1.pdf) [accessed Nov. 19, 2010].
- EPA (U.S. Environmental Protection Agency). 2002c. Determination of the Appropriate FQPA Safety Factor(s) in Tolerance Assessment. Office of Pesticide Programs, U.S. Environmental Protection Agency, Washington, DC. February 28, 2002 [online]. Available: <http://www.epa.gov/oppead1/trac/science/determ.pdf> [accessed Mar. 17, 2011].
- EPA (U.S. Environmental Protection Agency). 2005b. Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency [online]. Available: <http://www.epa.gov/osa/mmoaframework/pdfs/CANCER-GUIDELINES-FINAL-3-25-05%5B1%5D.pdf> [accessed Nov. 19, 2010].
- EPA (U.S. Environmental Protection Agency). 2005c. Supplemental Guidance for Assessing Cancer Susceptibility from Early-Life Exposure to Carcinogens. EPA/630/R-03/003F. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. March 2005 [online]. Available: [http://www.epa.gov/ttn/atw/childrens\\_supplement\\_final.pdf](http://www.epa.gov/ttn/atw/childrens_supplement_final.pdf) [accessed Mar. 17, 2011].
- EPA (U.S. Environmental Protection Agency). 2006. A Framework for Assessing Health Risks of Environmental Exposures to Children. U.S. Environmental Protection Agency [online]. Available: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=158363> [accessed Nov. 19, 2010].
- Gray, G.M., S.I. Baskin, G. Charnley, J.T. Cohen, L.S. Gold, N.I. Kerkvliet, H.M. Koenig, S.C. Lewis, R.M. McClain, L.R. Rhomberg, J.W. Snyder, and L.B. Weekley. 2001. The Annapolis accords on the use of toxicology in risk assessment and decision-making: An Annapolis Center Workshop report. *Toxicol. Mech. Methods* 11(3):225-231.

## **APPENDIX C – 6**

DRAFT—DO NOT CITE OR QUOTE

EPA/635/R-11/002A  
www.epa.gov/iris



# **TOXICOLOGICAL REVIEW**

## **OF**

# **LIBBY AMPHIBOLE ASBESTOS**

**In Support of Summary Information on the  
Integrated Risk Information System (IRIS)**

*August 2011*

*(Note: This document is an assessment of the noncancer and cancer health effects  
associated with the inhalation route of exposure only)*

### **NOTICE**

This document is an **External Review draft**. This information is distributed solely for the purpose of predissemination peer review under applicable information quality guidelines. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent any Agency determination or policy. It is being circulated for review of its technical accuracy and science policy implications.

U.S. Environmental Protection Agency  
Washington, DC

## 1. INTRODUCTION

This document presents background information and justification for the Integrated Risk Information System (IRIS) Summary of the hazard and exposure-response assessment of Libby Amphibole asbestos,<sup>1</sup> a mixture of amphibole fibers identified in the Rainy Creek complex and present in ore from the vermiculite mine near Libby, MT. IRIS Summaries may include oral reference dose (RfD) and inhalation reference concentration (RfC) values for chronic and other exposure durations, and a carcinogenicity assessment. This assessment reviews the potential hazards, both cancer and noncancer health effects, from exposure to Libby Amphibole asbestos and provides quantitative information for use in risk assessments: an RfC for noncancer and an inhalation unit risk addressing cancer risk. Libby Amphibole asbestos-specific data are not available to support RfD or cancer slope factor derivations for oral exposures.

An RfC is typically defined as “an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.” In the case of Libby Amphibole asbestos, the RfC is expressed in terms of the lifetime exposure in units of fibers per cubic centimeter of air (fibers/cc) in units of the fibers as measured by phase contrast microscopy (PCM). The inhalation RfC for Libby Amphibole asbestos considers toxic effects for both the respiratory system (portal-of-entry) and for effects peripheral to the respiratory system (extrarespiratory or systemic effects) that may arise after inhalation of Libby Amphibole asbestos. In this assessment, the estimates of hazard are derived from modeling cumulative exposures from human data, and thus for exposures of less than a lifetime the risk assessor should calculate a lifetime average concentration to compare to the RfC.

The carcinogenicity assessment provides information on the carcinogenic hazard potential of the substance in question, and quantitative estimates of risk from inhalation exposures are derived. The information includes a weight-of-evidence judgment of the likelihood that the agent is a human carcinogen and the conditions under which the carcinogenic effects may be expressed. Quantitative risk estimates are derived from the application of a low-

---

<sup>1</sup> The term “Libby Amphibole asbestos” is used in this document to identify the mixture of amphibole mineral fibers of varying elemental composition (e.g., winchite, richterite, tremolite, etc.) that have been identified in the Rainy Creek complex near Libby, MT. It is further described in Section 2.2.

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 dose extrapolation procedure from human data. An inhalation unit risk (IUR) is typically  
2 defined as a plausible upper bound on the estimate of cancer risk per  $\mu\text{g}/\text{m}^3$  air breathed for  
3 70 years. For Libby Amphibole asbestos, the RfC is expressed as a Lifetime Daily Exposure in  
4 fibers/cc (in units of the fibers as measured by PCM), and the IUR is expressed as cancer risk per  
5 fibers/cc (in units of the fibers as measured by PCM).

6 Development of these hazard identification and exposure-response assessments for Libby  
7 Amphibole asbestos has followed the general guidelines for risk assessment as set forth by the  
8 National Research Council (1983). U.S. Environmental Protection Agency (EPA) Guidelines  
9 and Risk Assessment Forum technical panel reports that may have been used in the development  
10 of this assessment include the following: *Guidelines for the Health Risk Assessment of Chemical*  
11 *Mixtures* (U.S. EPA, 1986c), *Guidelines for Mutagenicity Risk Assessment* (U.S. EPA, 1986b),  
12 *Recommendations for and Documentation of Biological Values for Use in Risk Assessment* (U.S.  
13 *EPA, 1988b*), *Guidelines for Developmental Toxicity Risk Assessment* (U.S. EPA, 1991a),  
14 *Interim Policy for Particle Size and Limit Concentration Issues in Inhalation Toxicity* (U.S. EPA,  
15 *1994a*), *Methods for Derivation of Inhalation Reference Concentrations and Application of*  
16 *Inhalation Dosimetry* (U.S. EPA, 1994b), *Use of the Benchmark Dose Approach in Health Risk*  
17 *Assessment* (U.S. EPA, 1995), *Guidelines for Reproductive Toxicity Risk Assessment* (U.S. EPA,  
18 *1996*), *Guidelines for Neurotoxicity Risk Assessment* (U.S. EPA, 1998), *Science Policy Council*  
19 *Handbook: Risk Characterization* (U.S. EPA, 2000c), *Benchmark Dose Technical Guidance*  
20 *Document* (U.S. EPA, 2000a), *Supplementary Guidance for Conducting Health Risk Assessment*  
21 *of Chemical Mixtures* (U.S. EPA, 2000d), *A Review of the Reference Dose and Reference*  
22 *Concentration Processes* (U.S. EPA, 2002), *Guidelines for Carcinogen Risk Assessment* (U.S.  
23 *EPA, 2005a*), *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to*  
24 *Carcinogens* (U.S. EPA, 2005b), *Science Policy Council Handbook: Peer Review* (U.S. EPA,  
25 *2006d*), and *A Framework for Assessing Health Risks of Environmental Exposures to Children*  
26 (U.S. EPA, 2006b).

27 The literature search strategy employed for this assessment is based on EPA's National  
28 Center for Environmental Assessment's Health and Environmental Research Outline database  
29 tool (which includes PubMed, MEDLINE, Web of Science, JSTOR, and other literature  
30 sources). The key search terms included the following: Libby Amphibole, tremolite, asbestos,  
31 richterite, winchite, amphibole, and Libby, MT. The relevant literature was reviewed through

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 July 2011. Any pertinent scientific information submitted by the public to the IRIS Submission  
2 Desk was also considered in the development of this document.  
3

#### 4 1.1. RELATED ASSESSMENTS

##### 5 1.1.1. IRIS Assessment for Asbestos (U.S. EPA, 1988a)

6 The IRIS assessment for asbestos was posted online in IRIS in 1988 and includes an IUR  
7 of 0.23 excess cancers per 1 fiber/cc (U.S. EPA, 1988a) (this unit risk is given in units of the  
8 fibers as measured by PCM). The IRIS IUR for general asbestos is derived by estimation of  
9 excess cancers for a continuous lifetime exposure and is based on the central tendency—not the  
10 upper bound—of the risk estimates (U.S. EPA, 1988a) and is applicable to exposures across a  
11 range of exposure environments and types of asbestos (CAS Number 1332-21-4). Although  
12 other cancers have been associated with asbestos (e.g., laryngeal, stomach, ovarian) (Straif et al.,  
13 2009), the IRIS IUR for asbestos accounts for only lung cancer and mesothelioma. Additionally,  
14 pleural and pulmonary effects from asbestos exposure (e.g., localized pleural thickening,  
15 asbestosis, and reduced lung function) are well documented, though, currently, there is no RfC  
16 for these noncancer health effects.

17 The derivation of the unit risk for general asbestos is based on the *Airborne Asbestos*  
18 *Health Assessment Update* (AAHAU) (U.S. EPA, 1986a). The AAHAU provides various cancer  
19 potency factors and mathematical models of lung cancer and mesothelioma mortality based on  
20 synthesis of data from occupational studies and presents estimates of lifetime cancer risk for  
21 continuous environmental exposures (0.0001 fiber/cc and 0.01 fiber/cc) (U.S. EPA, 1986a) (see  
22 Table 6-3). For both lung cancer and mesothelioma, life-table analysis was used to generate risk  
23 estimates based on the number of years of exposure and the age at onset of exposure. Although  
24 various exposure scenarios were presented, the unit risk is based on a lifetime continuous  
25 exposure from birth. The final asbestos IUR is 0.23 excess cancer per 1 fiber/cc continuous  
26 exposure<sup>2</sup> and was established by the EPA Carcinogen Risk Assessment Verification Endeavor  
27 workgroup and posted on the IRIS database in 1988 (U.S. EPA, 1988a) (see Table 1-1).  
28

---

<sup>2</sup>An IUR of 0.23 can be interpreted as a 23% increase in lifetime risk of dying from mesothelioma or lung cancer with each 1 fiber/cc increase in continuous lifetime exposure.

*This document is a draft for review purposes only and does not constitute Agency policy.*

Table 1-1. Derivation of the current IRIS inhalation unit risk for asbestos from the lifetime risk tables in the AAHAU

Gender	Excess deaths per 100,000 <sup>a</sup>			Risk	Unit risk
	Mesothelioma	Lung cancer	Total		
Female	183	35	218.5	$2.18 \times 10$	
Male	129	114	242.2	$2.42 \times 10$	
All	156	74	230.3	$2.30 \times 10$	0.23

<sup>a</sup>Data are for exposure at 0.01 fibers/cc for a lifetime.  
AAHAU = Airborne Asbestos Health Assessment Update.  
Source: U.S. EPA (1988a).

#### 1.1.2. EPA Health Assessment for Vermiculite (1991b)

An EPA health assessment for vermiculite reviewed available health data, including studies on workers who mined and processed ore with no significant amphibole fiber content. The cancer and noncancer health effects observed in the Libby, MT worker cohort were not seen in studies of workers exposed to vermiculite from mines with similar exposure to vermiculite but much lower exposures to asbestos fibers. Therefore, it was concluded that the health effects observed from the materials mined from Zonolite Mountain near Libby, MT, were most likely due to amphibole fibers not the vermiculite itself (U.S. EPA, 1991b). At the time, EPA recommended the application of the IRIS IUR for asbestos fibers (0.23 per fiber/cc) in addressing potential risk of the amphibole fibers entrained in vermiculite mined in Libby, MT.

#### 1.2. LIBBY AMPHIBOLE ASBESTOS-SPECIFIC HUMAN HEALTH ASSESSMENT

Libby Amphibole asbestos is a complex mixture of amphibole fibers—both mineralogically and morphologically (see Section 2.2). The mixture primarily includes tremolite, winchite, and richterite fibers with trace amounts of magnesioriebeckite, edenite, and magnesio-arfvedsonite. These fibers exhibit a complete range of morphologies from prismatic crystals to asbestiform fibers (Meeker et al., 2003). Epidemiologic studies of workers exposed to Libby Amphibole asbestos fibers indicate increased lung cancer and mesothelioma, as well as asbestosis, and other nonmalignant respiratory diseases (Larson et al., 2010b; Larson et al., 2010a; Moolgavkar et al., 2010; Rohs et al., 2008; Sullivan, 2007; McDonald et al., 2004, 2002;

*This document is a draft for review purposes only and does not constitute Agency policy.*

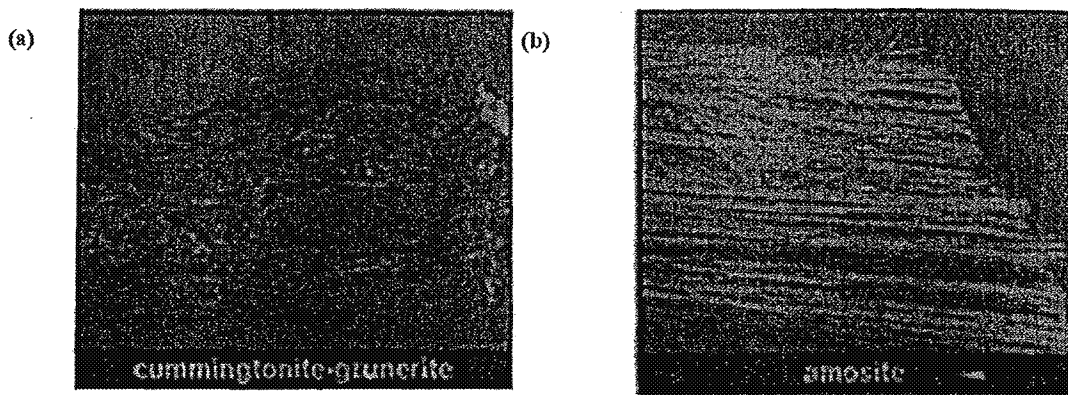


Figure 2-6. Comparison of crystalline forms amphibole minerals. Panel A shows a specimen identified as an amphibole mineral in the cummingtonite-grunerite solid solution series, although crystalline in form, the habit of formation did not favor formation of individual particles and fibers, hence its appearance as 'massive'. Panel B shows an amphibole mineral with very similar elemental composition but formed in a habit where very long fibers were allowed to form—hence the asbestiform appearance.

Source: Adapted from Bailey (2006).

may be elongated, but differ from the crystals described above as at least one face of the structure is the cleavage plane—not the face of a formed crystal.

With respect to classifying mineral field samples, geologists applied descriptive terms appropriate for viewing samples simply or at low magnification (e.g., field glass). The geologic terms for fiber morphology for classification of field samples is based on the macroscopic appearance of the crystals and fibers (e.g., acicular "needle-like in form") (AGL, 2005). In this framework, asbestos and asbestiform fibers are defined as long, slender, hair-like fibers visible to the naked eye (see Figure 2-6). This is a hallmark of commercially mined asbestos which is sought after for numerous applications because of its high tensile strength, heat resistance and in some cases, can be woven. Although these terms were used to describe fibers in hand samples and identify commercially valuable asbestos they are only applicable at the macroscopic level. It is important to realize that material defined as commercial asbestos, mined, milled, and manufactured into products not only contained these visible fibers, but many smaller fibers and single crystals which were not visible to the naked eye (Dement and Harris, 1979). As further

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 explained in Section 3, only these smaller fibers can enter the lung and transport to the pleura  
2 where the health effects of asbestos are best characterized. Therefore, for the purposes of this  
3 assessment (i.e., examining the health effects of asbestos fibers), consideration must be given to  
4 how these microscopic fibers are defined. For this purpose, terms intended for describing field  
5 samples may need to set aside, or redefined when applied at the microscopic level.

6 Currently there are several technologies commonly used to view and identify mineral  
7 structures at high magnification using light microscopes or electron microscopy. As standard  
8 analytical methods were developed for counting mineral fibers, structures and matrices using  
9 these instruments, analytical definitions to describe fibers and structures were developed. Phase  
10 contrast microscopy (PCM) was developed to detect fibers in occupational settings and has been  
11 widely used to assess worker exposure (see Text Box 2-1). The definition of a PCM-fiber is  
12 based purely on its dimensions. The standardization of the PCM method (i.e., NIOSH 7400) and  
13 its importance in applying health standards in occupational settings, results the common usage of  
14 the term 'fiber' to refer to those objects counted in the PCM analytical method (NIOSH, 1994a).  
15 However, this method cannot define the material or morphology of the viewed fiber. Thus  
16 PCM-fibers may be any material, and if they are mineral  
17 fibers may be any fiber morphology. If the nature of the  
18 fiber needs to be defined, NIOSH Method 7402 employs  
19 electron microscopy to determine if the fibers viewed by  
20 PCM are mineral fibers, and can establish the mineral  
21 composition (NIOSH, 1994b). This method does not  
22 recount the fibers, but, rather, it identifies what proportion  
23 of the fibers are mineral fibers, with an elemental  
24 composition consistent with asbestos, which is then used  
25 to adjust the PCM-fiber count. Although the PCM-fiber  
26 definition was not based on either mineralogy or an  
27 understanding of which fibers might be biologically  
28 relevant, this definition has become the basis of existing  
29 health standards (e.g., MSHA, 2008; OSHA, 1994; U.S.  
30 EPA, 1988a).

**Text Box 2-1: Fibers Viewed by Light Microscopy**

The collection of fibers on an air filter, and visually counted under a phase contrast microscope (PCM), was first described in 1934 by the Dutch physicist, Jits Zernike. The specification of a fiber as  $> 5 \mu\text{m}$  in length and length to diameter ratio (i.e., aspect ratio) of at least 3:1 resulted from this method. As a light microscope technique, the PCM method cannot distinguish mineral fibers from other fibers.

The U.S. Public Health Service developed and tested a standard air sampling method based on PCM detection (i.e., National Institute for Occupational Safety and Health [NIOSH] Method 7400). The NIOSH method specifies the analyst count fibers  $> 5 \mu\text{m}$  in length with an aspect ratio of at least 3:1. Results from PCM analysis are reported as fibers per cubic centimeter of air (f/cc) (See).

Electron microscopy can view objects at much higher magnification and can be coupled with other techniques which can identify the mineralogy (see Text Box 2-2). X-ray diffraction (XRD) may be used with the above techniques to differentiate crystalline structure of minerals in solid materials and provides information on the availability of the total mineral present. Thus, XRD can determine the mineral composition of the material analyzed, identifying its solid solution series and classifying the mineral per standardized nomenclature for amphibole minerals (see Section 2.1.1.1).

With the advent of the use of electron microscopy to identify mineral particles, there has been an attempt to resolve the traditional dimensional fiber definition(s), by describing the particles examined by electron microscopy and

X-ray diffraction in terms that are both geologically and mineralogically relevant. Structures viewed by electron microscopy may be described as having parallel sides, and considered 'fibers'. Where long, thin, curving fibers are viewed they may be described as 'asbestiform'. Structures with nonparallel sides can be considered acicular or prismatic, depending on their proportions. Thus, the descriptive terms used by geologists have migrated into the analytical field. However, the habit of formation of a single structure viewed by electron microscopy cannot be determined, and, while descriptive, these terms may not correlate to the geologic and commercial definitions of these terms. Therefore, the use of these definitions to describe individual particles viewed by TEM can be problematic (Meeker et

al., 2003). Important characteristics such as crystal structure and surface chemistry cannot be adequately categorized solely with visually determined definitions developed for the classification of field samples.

#### Text Box 2-2: Minerals Viewed by Electron Microscopy

Electron microscopy employs electrons, rather than light, to visualize the specimen. Furthermore, instead of using glass lenses to focus the light (wavelengths), electromagnetic lenses are used to focus electrons on the sample. The analytical techniques included in electron microscopy for asbestos testing are TEM (transmission electron microscopy), SEM (scanning electron microscopy), and STEM (scanning transmission electron microscopy). TEM produces two-dimensional (2-D) images that generally use a magnification factor of about 500 to 500,000 $\times$ . SEM produces three-dimensional (3-D) images that generally result in about 10 to 300,000 $\times$  magnification. STEM can produce both 2-D and 3-D images that generally result in about 10 to 500,000 $\times$  magnification.

The ISO 10312 method for analyzing air filters enumerates structures much smaller than the PCM fibers, with a minimum length requirement of 0.5  $\mu$ m. Additionally, structures with an aspect ratio of at least 5:1 are considered fibers, rather than 3:1, as with PCM analysis. The ISO 10312 method also defines other structures (rod bundles, clusters, and matrices) that are included in the structure count. Therefore, the term 'structure' rather than 'fiber' is used when presenting air sampling results from the ISO 10312 method where structures per cc of air (s/cc) are reported.

*This document is a draft for review purposes only and does not constitute Agency policy.*

### 3. FIBER TOXICOKINETICS

There are no published data on the toxicokinetics of Libby Amphibole asbestos.<sup>8</sup> However, to help inform the reader as to the expected toxicokinetics of Libby Amphibole asbestos, this section contains a general summary description of toxicokinetics of fibers. A more detailed discussion of fiber toxicokinetics is beyond the scope of this document and is reviewed elsewhere (NIOSH, 2011; ICRP, 1994).

The principal components of fiber toxicokinetics in mammalian systems are (1) deposition at the lung epithelial surface, and (2) clearance from the lung due to physical and biological mechanisms (including both translocation from the lung to other tissues [including the pleura]), and elimination from the body (see Figure 3-1).

Libby Amphibole asbestos includes fibers with a range of mineral compositions including amphibole fibers primarily identified as richterite, winchite, and tremolite (see Section 2.2). Although the fiber size varies somewhat from sample to sample, a large percentage (~45%) is less than 5  $\mu\text{m}$  long in bulk samples examined from the Libby mine site (Meeker et al., 2003). Limited data from air samples taken in the workplace also document a large percentage of fibers (including both respirable<sup>9</sup> fibers as well as fibers <5  $\mu\text{m}$ -long) (see Section 4.1.1.2 and Table 4-3). The importance of the size of fibers and how they deposit following inhalation is described below. Due to a lack of data specific to Libby Amphibole asbestos, these deposition steps are discussed for general forms of asbestos. The main route of human exposure to mineral fibers is through inhalation, although other routes of exposure play a role. Exposure of pulmonary tissue to fibers via the inhalation route depends on the fiber concentration in the breathing zone, the physical (aerodynamic) characteristics of the fibers, and the anatomy and physiology of the respiratory tract. Ingestion is another pathway of human exposure and occurs mainly through the swallowing of material removed from the lungs via mucociliary clearance or drinking water contaminated with asbestos, or eating, drinking, or smoking in asbestos-contaminated work environments (Condie, 1983). Handling asbestos can result in

<sup>8</sup>The term "Libby Amphibole asbestos" is used in this document to identify the mixture of amphibole mineral fibers of varying elemental composition (e.g., winchite, richterite, tremolite, etc.) that have been identified in the Rainy Creek complex near Libby, MT. It is further described in Section 2.2.

<sup>9</sup>Respirable fibers are those that can be inhaled into the lower lung where gas exchange occurs and are defined by their aerodynamic diameter ( $d_a \leq 3 \mu\text{m}$ ; NIOSH) (2011).

*This document is a draft for review purposes only and does not constitute Agency policy.*

#### 4. HAZARD IDENTIFICATION OF LIBBY AMPHIBOLE ASBESTOS

Several human studies are available that provide evidence for the hazard identification of Libby Amphibole asbestos.<sup>11</sup> This discussion focuses primarily on data derived from studies of people exposed to Libby Amphibole asbestos—either at work or in the community. The adverse health effects in humans are supported by the available Libby Amphibole asbestos experimental animal and laboratory studies. Libby Amphibole asbestos contains winchite (84%), with lesser amounts of richterite (11%) and tremolite (6%) with trace amounts of magnesiorichterite, edenite, and magnesio-arfvedsonite (Meeker et al., 2003) (see Section 2.2.3 for a more complete discussion). Adverse health effects from tremolite exposure have been reported in both human communities and laboratory animals; these effects are consistent with the human health effects reported for Libby Amphibole asbestos. Studies examining the health effects of exposure to winchite or richterite alone were not available in the published literature. The presentation of noncancer and cancer health effects provides a comprehensive review of adverse health effects observed from exposures to Libby Amphibole asbestos.

##### 4.1. STUDIES IN HUMANS—EPIDEMIOLOGY

The Libby Amphibole asbestos epidemiologic database includes studies conducted in occupational settings examining exposures to workers and community-based studies, which can include exposures to workers, exposures to family members of workers, and exposures from environmental sources. Occupational epidemiology studies exist for two worksites where workers were exposed to Libby Amphibole asbestos. These worksites include the mine and mill at the Zonolite Mountain operations near Libby, MT, and a vermiculite processing plant in Marysville, OH. Worker cohorts from each site and the study results are described in Section 4.1.1. Community-based studies include community health consultations for Libby, MT conducted by the Agency for Toxic Substances and Disease Registry (ATSDR), including an evaluation of cancer mortality data, and a health screening of current and former area residents—including workers—that collected medical and exposure histories, chest X-rays, and pulmonary function tests (ATSDR, 2001b, 2000) (see Section 4.1.2). ATSDR, in conjunction

<sup>11</sup> The term “Libby Amphibole asbestos” is used in this document to identify the mixture of amphibole mineral fibers of varying elemental composition (e.g., winchite, richterite, tremolite, etc.) that have been identified in the Rainy Creek complex near Libby, MT. It is further described in Section 2.2.

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 with state health departments, also conducted health consultations for 28 other communities  
2 around vermiculite processing plants that were potentially exposed to Libby Amphibole asbestos  
3 (see Section 4.1.4). These health consultations consisted of analyses of cancer incidence or  
4 mortality data; results from nine of these studies are currently available.

5 No occupational studies are available for exposure to tremolite, richterite, or winchite  
6 mineral fibers individually or as a mixture exposure, other than Libby Amphibole asbestos.  
7 Communities, however, have been exposed to tremolite and other mineral fibers from natural  
8 soils and outcroppings. Tremolite asbestos-containing soil has been used in whitewash in  
9 interior wall coatings in parts of Turkey and Greece. Studies in these areas published as early as  
10 1979 reported an increased risk of pleural and peritoneal malignant mesothelioma (Sichletidis et  
11 al., 1992; Baris et al., 1987; Langer et al., 1987; Baris et al., 1979). More recent studies of  
12 communities exposed to tremolite and chrysotile fibers report excess lung cancer and  
13 mesothelioma (1.3- and 6.9-fold, respectively) (Hasanoglu et al., 2006). Other studies reported  
14 pleural anomalies in residents exposed to naturally occurring asbestos, which includes actinolite,  
15 tremolite, and anthophyllite (Metintas et al., 2005; Zeren et al., 2000). Clinical observations  
16 include a bilateral increase in pleural calcification accompanied by restrictive lung function as  
17 the disease progresses, a condition known as “Metsovo lung,” named after a town in Greece  
18 (Constantopoulos et al., 1985). In one community, the prevalence of pleural calcification was  
19 46% (of 268 residents), increasing with age to 80% in residents over 70 (Langer et al., 1987).  
20 Both tremolite and chrysotile were identified in bronchoalveolar lavage fluid of 65 residents  
21 from different areas of Turkey who were environmentally exposed (Dumortier et al., 1998). The  
22 health effects observed in communities with environmental and residential exposure to tremolite  
23 are consistent with health effects documented for workers exposed to commercial forms of  
24 asbestos.

#### 25 26 **4.1.1. Studies of Libby, MT Vermiculite Mining Operation Workers**

27 Several studies of mortality from specific diseases among workers in the Libby, MT  
28 mining operations have been conducted, beginning in the 1980s with the studies by McDonald  
29 et al. (1986a) and Amandus and Wheeler. (1987). McDonald et al. (2004, 2002) published an  
30 update with mortality data through 1999, and Sullivan (2007) updated the cohort originally  
31 described by Amandus and Wheeler (1987) (referred to in this assessment as the Libby worker

*This document is a draft for review purposes only and does not constitute Agency policy.*

1       Larson et al. (2010b) evaluated multiple causes of death, and, therefore, more than one  
2       cause of death can be coded for an individual. A total of 104 lung or bronchus cancer deaths  
3       were observed, for an SMR of 1.6 (95% CI: 1.3, 2.0) using an external comparison of United  
4       States cause of death data from 1960 to 2002 (Larson et al., 2010b). A higher risk was seen in  
5       the higher cumulative exposure categories using Cox-proportional hazards modeling with an  
6       internal referent group: relative risk 1.0 (referent), 1.1 (95% CI: 0.6, 2.1), 1.7 (95% CI: 1.0, 3.0),  
7       and 3.2 (95% CI: 1.8, 5.3) respectively, for <1.4 (referent), 1.4 to <8.6, 8.6 to <44.0 and ≥44.0  
8       fibers/cc-years. Larson et al. (2010b) used data from a health screening program conducted in  
9       Libby by ATSDR in 2000–2001 (described in Section 4.1.2.2) pertaining to smoking history to  
10      estimate that the proportion of smokers ranged from 50% to 66% in the unexposed group  
11      (defined as exposure <8.6 fibers/cc-years) and between 66% and 85% among the exposed  
12      (defined as ≥8.6 fibers/cc-years). Larson et al. (2010b) used these estimates in a Monte Carlo  
13      simulation to estimate the potential bias in lung cancer risks that could have been introduced by  
14      differences in smoking patterns. The bias-adjustment factor ( $RR_{unadjusted}/RR_{adjusted} = 1.3$ ) reduced  
15      the overall RR estimate for lung cancer from 2.4 to 2.0.

#### 17   4.1.1.3.2. *Mesothelioma*

18      Data pertaining to mesothelioma risk from the available studies are summarized in  
19      Table 4-5. McDonald et al. (2004) presented dose-response modeling of mesothelioma risk  
20      based on 12 cases. Using Poisson regression, the mesothelioma mortality rate across increasing  
21      categories of exposure was compared to the rate in the lowest exposure category. Note that the  
22      referent group was also at excess risk of dying from mesothelioma; that is, one to three cases of  
23      mesothelioma were observed in the referent group, depending on the exposure index. Three  
24      exposure indices were used in analysis: average intensity over the first 5 years of employment,  
25      cumulative exposure, and residence-weighted cumulative exposure. Because of the requirement  
26      for 5 years of employment data, 199 individuals (including three mesothelioma cases) were  
27      excluded from the analysis of average intensity. The residence-weighted cumulative exposure  
28      was based on the summation of exposure by year, weighted by years since the exposure. This  
29      metric gives greater weight to exposures that occurred a longer time ago. Although evidence of  
30      an excess risk of dying from mesothelioma was seen in all groups, there was little evidence of  
31      increasing RR with increasing average intensity or cumulative exposure. For the

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 al., 2010b; Sullivan, 2007; McDonald et al., 2004)<sup>16</sup> observed increasing risks with increasing  
2 cumulative exposure exposures when analyzed using tertiles or quartiles, or as a continuous  
3 measure. Increased risks are also seen in the studies reporting analyses using an external referent  
4 group, i.e., standardized mortality ratios (Sullivan, 2007; Amandus and Wheeler, 1987;  
5 McDonald et al., 1986a). Radiographic evidence of small opacities (evidence of parenchymal  
6 damage) and pleural thickening (both discrete and diffuse) has also been shown in studies of  
7 Libby workers (Larson et al., 2010a; Whitehouse, 2004; Amandus et al., 1987b; McDonald et al.,  
8 1986b).

#### 10 4.1.2. Libby, MT Community Studies

11 In addition to worker exposures, the operations of the Zonolite Mountain mine are  
12 believed to have resulted in both home exposures and community exposures. Potential pathways  
13 of exposure (discussed below) range from release of airborne fibers into the community,  
14 take-home exposure from mine workers (e.g., clothing), and recreational activities including  
15 gardening and childhood play activities. Due to a potential for a broader community concern,  
16 ATSDR conducted several studies and health actions responding to potential asbestos  
17 contamination in the Libby, MT area.

##### 19 4.1.2.1. Geographic Mortality Analysis

20 ATSDR conducted a location-specific analysis of mortality risks and a community health  
21 screening for asbestos in the Libby area (see Table 4-8). The mortality analysis was based on  
22 death certificate data from 1979–1998, with geocoding of current residence at time of death. The  
23 six geographic areas used in the analysis were defined as the Libby city limits (1.1 square miles  
24 around the downtown); the extended boundary of Libby (2.2 square miles around the  
25 downtown); the boundary based on air modeling (16 square miles, based on computer modeling  
26 of asbestos fiber distribution); the medical screening boundary (25 square miles, including the  
27 town of Libby and areas along the Kootenai River); the Libby valley (65 square miles); and  
28 central Lincoln County (314 square miles, based on a 10-mile radius around downtown Libby)  
29 (ATSDR, 2000).

---

<sup>16</sup>See also reanalysis of Sullivan (2007) data by Moolgavar et al. (2010).

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 The 1990 population estimates were 2,531, 3,694, 4,300, 6,072, 8,617, and 9,512,  
2 respectively, for these six areas. Age-standardized SMRs were calculated using underlying  
3 cause-of-death information obtained from death certificates issued during the study period for  
4 413 of 419 identified decedents, and Montana and U.S. populations were used as reference  
5 groups. Increased SMRs were observed for both asbestosis and pulmonary circulation diseases  
6 (see Table 4-8). The SMR for lung cancer ranged from 0.9–1.1 and 0.8–1.0 in the analyses for  
7 each of the six geographic boundaries using Montana and U.S. reference rates, respectively. In  
8 addition, four deaths due to mesothelioma were observed during the study period. These  
9 analyses did not distinguish between deaths among workers and deaths among other community  
10 members.

#### 12 4.1.2.2. Community Screening—Respiratory Health

13 The ATSDR community health screening was conducted from July–November 2000 and  
14 July–September 2001 with 7,307 total participants (ATSDR, 2001b) (see Table 4-9). Eligibility  
15 was based on residence, work, or other presence in Libby for at least 6 months before 1991. The  
16 total population eligible for screening is not known; the population of Libby, MT in 2000 was  
17 approximately 10,000. In addition to a standardized interview regarding medical history,  
18 symptoms, work history, and other potential exposures, clinical tests included spirometry (forced  
19 expiratory volume in one second [FEV1] and FVC) and chest X-rays (for participants aged  
20 18 years and older). Moderate to severe restriction (defined by the researchers as FVC <70%  
21 predicted value) was observed in 2.2% of the men and 1.6% of women but was not observed in  
22 individuals less than age 18.

23 Two board-certified radiologists (B readers) examined each radiograph, and a third reader  
24 was used in cases of disagreement. Readers were aware that the radiographs were from  
25 participants in the Libby, MT health screening but were not made aware of exposure histories  
26 and other characteristics (Peipins et al., 2004a; Price, 2004; Peipins et al., 2003). The  
27 radiographs revealed pleural abnormalities in 17.9% of participants, with prevalence increasing  
28 with increasing number of “exposure pathways” (defined on the basis of potential work and  
29 residential exposure to asbestos within Libby and from other sources) (see Table 4-9). Detailed  
30 results of an analysis excluding the former Libby workers cohort were not presented, but the  
31 authors noted that the relationship between number of exposure pathways and increasing

*This document is a draft for review purposes only and does not constitute Agency policy.*

Table 4-8. Cancer mortality and nonmalignant respiratory disease mortality in the Libby, MT community

Reference(s)	Inclusion criteria and design details	Results
ATSDR (2000)	<p>1979–1998, underlying cause of death from death certificates; geocoding of street locations (residence at time of death) within six geographic boundaries (ranging from 2,532 residents in Libby city limits to 9,521 in central Lincoln County in 1990). Inquiries to postmaster were required because of P.O. Box address for 8% (<math>n = 32</math>); information on 47 of 91 residents of elderly care facilities resulted in reclassification of 16 of 47 (34%) to nonresidents of Libby.</p> <p>U.S. Census data corresponding to the same six geographic boundaries of Libby, MT.</p> <p>419 decedents identified, 418 death certificates obtained, 413 with geocoding.</p> <p>Age-standardized SMRs based on Montana and U.S. comparison rates. Asbestosis SMRs were somewhat higher using the U.S. referent group, but choice of referent group had little difference on SMRs for most diseases.</p> <p>Four deaths from mesothelioma observed in the study area.</p>	<p>Lung cancer (<math>n = 82</math>) SMR (95% CI)</p> <p>Comparison area (Montana reference rates):</p> <p>Libby city limits 1.1 (0.8, 1.5)</p> <p>Extended Libby boundary 1.1 (0.8, 1.5)</p> <p>Air modeling 1.0 (0.8, 1.4)</p> <p>Medical screening 0.9 (0.7, 1.2)</p> <p>Libby valley 0.9 (0.7, 1.2)</p> <p>Central Lincoln County 0.9 (0.7, 1.1)</p> <p>Pancreatic cancer (<math>n = 10</math>) SMR (95% CI)</p> <p>Comparison area (Montana reference rates):</p> <p>Libby city limits 1.0 (0.5, 2.1)</p> <p>Extended Libby boundary 0.9 (0.4, 1.7)</p> <p>Air modeling 0.7 (0.3, 1.4)</p> <p>Medical screening 0.7 (0.3, 1.2)</p> <p>Libby valley 0.6 (0.3, 1.0)</p> <p>Central Lincoln County 0.5 (0.3, 1.0)</p> <p>Asbestosis (<math>n = 11</math>) SMR (95% CI)</p> <p>Comparison area (Montana reference rates):</p> <p>Libby city limits 40.8 (13.2, 95.3)</p> <p>Extended Libby boundary 47.3 (18.9, 97.5)</p> <p>Air modeling 44.3 (19.1, 87.2)</p> <p>Medical screening 40.6 (18.5, 77.1)</p> <p>Libby valley 38.7 (19.3, 69.2)</p> <p>Central Lincoln County 36.3 (18.1, 64.9)</p> <p>Comparison area (U.S. reference rates):</p> <p>Libby city limits 63.5 (20.5, 148)</p> <p>Extended Libby boundary 74.9 (30.0, 154)</p> <p>Air modeling 71.0 (30.6, 140)</p> <p>Medical screening 66.1 (30.2, 125)</p> <p>Libby valley 63.7 (31.7, 114)</p> <p>Central Lincoln County 59.8 (29.8, 107)</p> <p>Pulmonary circulation (<math>n = 14</math>) SMR (95% CI)</p> <p>Comparison area (Montana reference rates):</p> <p>Libby city limits 2.3 (1.1, 4.4)</p> <p>Extended Libby boundary 1.9 (0.9, 3.7)</p> <p>Air modeling 1.8 (0.9, 3.3)</p> <p>Medical screening 1.6 (0.8, 2.9)</p> <p>Libby valley 1.6 (0.9, 2.7)</p> <p>Central Lincoln County 1.5 (0.8, 2.5)</p>

This document is a draft for review purposes only and does not constitute Agency policy.

Table 4-9. Pulmonary function and chest radiographic studies in the Libby, MT community

Reference(s)	Inclusion criteria and design details	Results																																																																																																				
Peipins et al. (2003); ATSDR (2001b)	Resided, worked, attended school, or participated in other activities in Libby for at least 6 months before 1991 (including mine employees and contractors). Health screening between July and November 2000. Conducted interviews ( <i>n</i> = 6,149, 60% of Libby residents based on 2000 Census data) and chest X-rays ( <i>n</i> = 5,590, 18 years and older), and determined spirometry—forced expiratory volume in 1 second (FEV1), forced vital capacity (FVC1), and ratio (FEV1/FVC). 19 “exposure pathways” including Libby mining company work, contractor work, dust exposure at other jobs, vermiculite exposure at other jobs, potential asbestos exposure at other jobs or in the military, cohabitation with Libby mining company worker, and residential and recreational use of vermiculite. Chest X-rays read by 1980 ILO classifications (3 views; posterior-anterior, right- and left- anterior oblique). Peipins et al. (2003) similar to (ATSDR, 2001b) except longer screening period (July–November 2000 and July–September 2001). Conducted interviews ( <i>n</i> = 7,307) and chest X-rays ( <i>n</i> = 6,668).	Peipins (2003) and ATSDR (2001b): Pleural abnormalities seen in 17.9% of participants; increasing prevalence with increasing number of exposure pathways (6.7% among those with no specific pathways, 34.6% among those with 12 or more pathways).  ATSDR (2001b): Moderate-to-severe FVC1 restriction (FVC <70% predicted): 2.2% of men >17 years old; 1.6% of women >17 years old; 0.0% of men or women <18 years old. Also includes data on self-reported lung diseases and symptoms.																																																																																																				
Weill et al. (2011)	Participants in the ATSDR community health screening (see first row in table). Analysis limited to ages 25 to 90 years, excluding individuals with history of other asbestos-related work exposures, with spirometry, consensus reading of chest X-ray, smoking data, and exposure pathway data ( <i>n</i> = 4,397). Analysis based on five exposure categories: (1) W.R. Grace worker, (2) other vermiculite worker (contractor work), (3) other dusty occupation, (4) household (combination of three household categories), and (5) environmental (“no” to work and household exposures in Categories 1–6). Chest X-rays read by 1980 ILO classifications (frontal view).	<table><thead><tr><th></th><th>Profusion ≥1/0</th><th>Plaque</th><th>DPT/ CAO</th></tr></thead><tbody><tr><td colspan="4">Prevalence (%), ages 25 to 40 years:</td></tr><tr><td>1) W.R. Grace</td><td>0.0</td><td>20.0</td><td>5.0</td></tr><tr><td>2) Other</td><td>0.8</td><td>0.8</td><td>0.0</td></tr><tr><td>3) Dusty</td><td>0.0</td><td>3.8</td><td>0.4</td></tr><tr><td>4) Household</td><td>0.0</td><td>2.2</td><td>0.0</td></tr><tr><td>5) Environment</td><td>0.0</td><td>0.4</td><td>0.0</td></tr><tr><td colspan="4">Prevalence (%), ages 41 to 50 years:</td></tr><tr><td>1) W.R. Grace</td><td>0.0</td><td>26.2</td><td>5.0</td></tr><tr><td>2) Other</td><td>0.5</td><td>7.8</td><td>1.0</td></tr><tr><td>3) Dusty</td><td>0.0</td><td>2.8</td><td>0.9</td></tr><tr><td>4) Household</td><td>0.0</td><td>11.1</td><td>0.4</td></tr><tr><td>5) Environment</td><td>0.0</td><td>1.9</td><td>0.2</td></tr><tr><td colspan="4">Prevalence (%), ages 51 to 60 years:</td></tr><tr><td>1) W.R. Grace</td><td>3.2</td><td>34.9</td><td>3.2</td></tr><tr><td>2) Other</td><td>0.6</td><td>13.7</td><td>0.6</td></tr><tr><td>3) Dusty</td><td>0.6</td><td>12.6</td><td>0.0</td></tr><tr><td>4) Household</td><td>1.0</td><td>20.1</td><td>1.5</td></tr><tr><td>5) Environment</td><td>0.0</td><td>7.7</td><td>0.9</td></tr><tr><td colspan="4">Prevalence (%), ages 61 to 90 years:</td></tr><tr><td>1) W.R. Grace</td><td>11.1</td><td>45.7</td><td>8.6</td></tr><tr><td>2) Other</td><td>0.6</td><td>24.8</td><td>8.5</td></tr><tr><td>3) Dusty</td><td>1.1</td><td>21.9</td><td>3.3</td></tr><tr><td>4) Household</td><td>2.4</td><td>38.3</td><td>5.7</td></tr><tr><td>5) Environment</td><td>1.3</td><td>12.7</td><td>2.2</td></tr></tbody></table>		Profusion ≥1/0	Plaque	DPT/ CAO	Prevalence (%), ages 25 to 40 years:				1) W.R. Grace	0.0	20.0	5.0	2) Other	0.8	0.8	0.0	3) Dusty	0.0	3.8	0.4	4) Household	0.0	2.2	0.0	5) Environment	0.0	0.4	0.0	Prevalence (%), ages 41 to 50 years:				1) W.R. Grace	0.0	26.2	5.0	2) Other	0.5	7.8	1.0	3) Dusty	0.0	2.8	0.9	4) Household	0.0	11.1	0.4	5) Environment	0.0	1.9	0.2	Prevalence (%), ages 51 to 60 years:				1) W.R. Grace	3.2	34.9	3.2	2) Other	0.6	13.7	0.6	3) Dusty	0.6	12.6	0.0	4) Household	1.0	20.1	1.5	5) Environment	0.0	7.7	0.9	Prevalence (%), ages 61 to 90 years:				1) W.R. Grace	11.1	45.7	8.6	2) Other	0.6	24.8	8.5	3) Dusty	1.1	21.9	3.3	4) Household	2.4	38.3	5.7	5) Environment	1.3	12.7	2.2
	Profusion ≥1/0	Plaque	DPT/ CAO																																																																																																			
Prevalence (%), ages 25 to 40 years:																																																																																																						
1) W.R. Grace	0.0	20.0	5.0																																																																																																			
2) Other	0.8	0.8	0.0																																																																																																			
3) Dusty	0.0	3.8	0.4																																																																																																			
4) Household	0.0	2.2	0.0																																																																																																			
5) Environment	0.0	0.4	0.0																																																																																																			
Prevalence (%), ages 41 to 50 years:																																																																																																						
1) W.R. Grace	0.0	26.2	5.0																																																																																																			
2) Other	0.5	7.8	1.0																																																																																																			
3) Dusty	0.0	2.8	0.9																																																																																																			
4) Household	0.0	11.1	0.4																																																																																																			
5) Environment	0.0	1.9	0.2																																																																																																			
Prevalence (%), ages 51 to 60 years:																																																																																																						
1) W.R. Grace	3.2	34.9	3.2																																																																																																			
2) Other	0.6	13.7	0.6																																																																																																			
3) Dusty	0.6	12.6	0.0																																																																																																			
4) Household	1.0	20.1	1.5																																																																																																			
5) Environment	0.0	7.7	0.9																																																																																																			
Prevalence (%), ages 61 to 90 years:																																																																																																						
1) W.R. Grace	11.1	45.7	8.6																																																																																																			
2) Other	0.6	24.8	8.5																																																																																																			
3) Dusty	1.1	21.9	3.3																																																																																																			
4) Household	2.4	38.3	5.7																																																																																																			
5) Environment	1.3	12.7	2.2																																																																																																			

This document is a draft for review purposes only and does not constitute Agency policy.

**Table 4-9. Pulmonary function and chest radiographic studies in the Libby, MT community (continued)**

Reference(s)	Inclusion criteria and design details	Results
Vinikoor et al. (2010)	Participants in the ATSDR community health screening (see first row in table). Analysis limited to $n = 1,003$ ages 10–29 years at time of health screening ( $\leq$ age 18 in 1990 when the mining/milling operations closed). Excluded if worked for W.R. Grace, or for a contractor of W.R. Grace, exposed to dust at other jobs, or exposed to vermiculite at other jobs. Exposure characterized by 6 activities (never, sometimes, or frequently participated in 1–2 or $\geq 3$ activities). Analysis of history of respiratory symptoms and spirometry data (obstructive, restrictive, or mixed).	Little difference across exposure levels in prevalence of physician-diagnosed lung disease or abnormal spirometry. Odds Ratio (95% CI) seen between $\geq 3$ activities and Usual cough 2.93 (0.93, 9.25) Shortness of breath 1.32 (0.51, 3.42) Bloody phlegm 1.49 (0.41, 5.43)

OR = odds ratio; DPT = diffuse pleural thickening; CAO = costophrenic angle obliteration.

prevalence of pleural abnormalities was somewhat attenuated with this exclusion. The prevalence of pleural anomalies decreased from approximately 35% to 30% in individuals with 12 or more exposure pathways when these workers were excluded from the analysis. Among individuals with no definable exposure pathways, the prevalence of pleural anomalies was 6.7%, which is higher than reported in other population studies (Peipins et al., 2004a; Price, 2004). The direct comparability between study estimates is difficult to make; the possibility of over- or underascertainment of findings from the X-rays based on knowledge of conditions in Libby was not assessed in this study. No information is provided regarding analyses excluding all potential work-related asbestos exposures.

Weill et al. (2011) used the ATSDR community health screening data to analyze the prevalence of X-ray abnormalities in relation to age, smoking history, and types of exposures. From the 6,668 participants with chest X-rays, 1,327 individuals with a history of asbestos-related work (other than with the Grace mining or related vermiculite operations) were excluded, along with 817 excluded based on age ( $<25$  or  $>90$  years) or lack of spirometric data, smoking data, or exposure pathway data. An additional 127 were excluded because a consensus agreement (2 out of 3 readers) was not reached regarding the X-ray findings, leaving  $n = 4,397$  in the analysis. Analysis was based on five exposure categories: (1) Grace worker ( $n = 255$ ), (2) other vermiculite worker (e.g., secondary contractor worker for Grace or other jobs with vermiculite exposure ( $n = 664$ ), (3) other dusty occupation (e.g., plumber, dry wall finisher,

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 carpenter, roofer, electrician, welder, shipyard work or ship construction or repair ( $n = 831$ ),  
2 (4) household, including household with other vermiculite or dusty work (lived with a Grace  
3 worker combination of three household categories) ( $n = 880$ ), and (5) environmental (“no” to  
4 work and household exposures in Categories 1–4) ( $n = 1,894$ ). The frontal views (posterior-  
5 anterior) of the chest X-rays were used in this analysis [in contrast to the use of frontal and  
6 oblique views in Peipins et al. (2003)]. As expected, lung function ( $FEV_1$ , FVC, and  $FEV_1/FVC$ )  
7 was lower among ever smokers compared with never smokers (within each age group) and  
8 decreased with age (within each smoking category). The prevalence of X-ray abnormalities  
9 (plaques, or diffuse pleural thickening, and/or costophrenic angle obliteration) also generally  
10 increased with age (divided into 25–40, 41–50, 51–60, and 61–90 years) within each of the  
11 exposure categories (see Table 4-9), with the highest prevalence seen among Grace workers. For  
12 a given age, the prevalence among those with environmental exposure only (i.e., no household or  
13 occupational exposures) was similar to the prevalence among those with non-Grace occupational  
14 or household exposures in the next youngest age category. The prevalence among the household  
15 contact category was similar or higher than the prevalence among the other vermiculite and dusty  
16 job categories. This household contact category includes individuals who lived with a Grace  
17 worker with no personal history of vermiculite or dust work ( $n = 594$ ) and those who also had a  
18 history of other vermiculite ( $n = 114$ ) or dusty ( $n = 172$ ) jobs. The authors noted the prevalence  
19 rates were similar among these groups, and so the analysis was based on the combination of  
20 these three groups. Mean FVCs ( $\pm$ SE) percentage predicted were 78.76 ( $\pm 3.64$ ), 82.16 ( $\pm 3.34$ ),  
21 95.63 ( $\pm 0.76$ ), and 103.15 ( $\pm 0.25$ ), respectively, in those with diffuse pleural thickening and/or  
22 costophrenic angle obliteration, profusion  $\geq 1/0$ , other pleural abnormalities, and no pleural  
23 abnormalities. The strongest effects of diffuse pleural thickening and/or costophrenic angle  
24 obliteration on FVC were seen among men who had never smoked ( $-23.77$ ,  $p < 0.05$ ), with  
25 smaller effects seen among men who had smoked ( $-9.77$ ,  $p < 0.05$ ) and women who had smoked  
26 ( $-6.73$ ,  $p < 0.05$ ).

27 Vinikoor et al. (2010) used the 2000–2001 health screening data to examine respiratory  
28 symptoms and spirometry results among 1,224 adolescents and young adults who were 18 years  
29 or younger in 1990 when the mining/milling operations closed. At the time of the health  
30 screening, the ages in this group ranged from 10 to 29 years. Exclusion criteria for this analysis  
31 included previous work for W.R. Grace, work for a contractor of W.R. Grace, exposure to dust at

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 other jobs, or exposure to vermiculite at other jobs. The total number of exclusions was 221,  
2 leaving 1,003 in the analysis. The potential for vermiculite exposure was classified based on  
3 responses to questions about six activities (handling vermiculite insulation, participation in  
4 recreational activities along the vermiculite-contaminated gravel road leading to the mine,  
5 playing at the ball fields near the expansion plant, playing in or around the vermiculite piles,  
6 heating the vermiculite to “pop” it, and other activities involving vermiculite). The medical  
7 history questionnaire included information on three respiratory symptoms: usually have a cough  
8 ( $n = 108$ , 10.8%); troubled by shortness of breath when walking up a slight hill or when hurrying  
9 on level ground ( $n = 145$ , 14.5%); coughed up phlegm that was bloody in the past year  
10 ( $n = 59$ , 5.9%). A question on history of physician-diagnosed lung disease ( $n = 51$ , 5.1%) was  
11 also included. The spirometry results were classified as normal in 896 (90.5%), obstructive in  
12 62 (6.3%), restrictive in 30 (3.0%), and mixed in 2 (0.2%). Information on smoking history was  
13 also collected in the questionnaire: 15.8% and 7.3% were classified as current and former  
14 smokers, respectively. Approximately half of the participants lived with someone who smoked.  
15 The analyses adjusted for age, sex, personal smoking history, and living with a smoker. For  
16 usually having a cough, the odds ratios (ORs) were 1.0 (referent), 1.88 (95% CI: 0.71, 5.00),  
17 2.00 (95% CI: 0.76, 5.28) and 2.93 (95% CI: 0.93, 9.25) for never, sometimes, frequently  
18 participated in 1–2 activities, and frequently participated in  $\geq 3$  activities, respectively. For  
19 shortness of breath, the corresponding ORs across those exposure categories were 1.0 (referent),  
20 1.16 (95% CI: 0.55, 2.44), 1.27 (95% CI: 0.61, 2.63) and 1.32 (95% CI: 0.51, 3.42), and for  
21 presence of bloody phlegm in the past year the ORs were 1.0 (referent), 0.85 (95% CI: 0.31,  
22 2.38), 1.09 (0.41, 2.98), and 1.49 (95% CI: 0.41, 5.43). For history of physician-diagnosed lung  
23 disease and abnormal spirometry results, there was little difference in the odds ratios across the  
24 exposure categories: for lung disease, the ORs were 1.0 (referent), 1.95 (95% CI: 0.57, 6.71),  
25 1.51 (95% CI: 0.43, 5.24) and 1.72 (95% CI: 0.36, 8.32) for the categories of never, sometimes,  
26 frequently participated in 1–2 activities, and frequently participated in  $\geq 3$  activities, respectively.  
27 For abnormal spirometry (i.e., obstructive, restrictive, or mixed,  $n = 94$  cases), the ORs were  
28 1.0 (referent), 1.34 (95% CI: 0.60, 2.96), 1.20 (95% CI: 0.53, 2.70) and 1.33 (95% CI: 0.42,  
29 4.19) across these exposure groups.

30 Two other studies examining autoimmune disease and autoantibodies in residents of  
31 Libby, Montana are described in Section 4.3.

*This document is a draft for review purposes only and does not constitute Agency policy.*

1    **4.1.2.3. Other Reports of Asbestos-Related Disease Among Libby, MT Residents**

2            Whitehouse et al. (2008) recently reviewed 11 cases of mesothelioma diagnosed between  
3    1993 and 2006 in residents in or around Libby, MT ( $n = 9$ ) and in family members of workers in  
4    the mining operations ( $n = 2$ ). Three cases were men who might have had occupational asbestos  
5    exposure through construction work (Case 1), working in the U.S. Coast Guard and as a  
6    carpenter (Case 5), or through railroad work involving sealing railcars in Libby (Case 7). One  
7    case was a woman whose father had worked at the mine for 2 years; although the family lived  
8    100 miles east of Libby, her exposure may have come through her work doing the family  
9    laundry, which included laundering her father's work clothes. The other seven cases  
10   (four women, three men) had lived or worked in Libby for 6–54 years, and had no known  
11   occupational or family-related exposure to asbestos. Medical records were obtained for all  
12   11 patients; pathology reports were obtained for 10 of the 11 patients. The Centers for Disease  
13   Control estimated the death rate from mesothelioma, using 1999 to 2005 data, as approximately  
14   14 per million per year (CDC, 2009), approximately five times higher than the rate estimated by  
15   Whitehouse et al. (2008) for the Libby area population based on the estimated population of  
16   9,500 for Lincoln County and 15 years (or 150,000 person-years) covered by the analysis.  
17   Whitehouse et al. (2008) stated that a W.R. Grace unpublished report of measures taken in 1975  
18   indicated that exposure levels of 1.1 fibers/cc were found in Libby, and 1.5 fibers/cc were found  
19   near the mill and railroad facilities. Because the mining and milling operations continued to  
20   1990, and because of the expected latency period for mesothelioma, Whitehouse et al. (2008)  
21   suggests that additional cases can be expected to occur within this population.

22  
23   **4.1.2.4. Summary of Respiratory Health Effects in Libby, MT Community Studies**

24            The geographic-based mortality analysis of 1997–1998 mortality data indicates that  
25   asbestos-related mortality is substantially increased in Libby, MT, and the surrounding area,  
26   with rates 40 times higher compared with Montana rates and 60–70 times higher compared with  
27   U.S. rates (ATSDR, 2000). These data provide evidence of the disease burden within the  
28   community; however, because this analysis did not distinguish between deaths among workers  
29   and deaths among other community members, it is not possible based on these data to estimate  
30   the risk of asbestos-related mortality experienced by residents who were not employed at the  
31   mining or milling operations. The community health screening studies provide more detailed

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 information regarding exposure pathways in addition to occupation (ATSDR, 2001b). Data from  
2 the ATSDR community health screening study indicate that the prevalence of pleural  
3 abnormalities, identified by radiographic examination, increases substantially with increasing  
4 number of exposure pathways (Peipins et al., 2003). In addition, the prevalence of some  
5 self-reported respiratory symptoms among 10 to 29-year-old adolescents and young adults was  
6 associated with certain exposure pathways. These participants were  $\leq$  age 18 in 1990 when the  
7 mining/milling operations closed (Vinikoor et al., 2010). A better understanding of the  
8 community health effects and the examination of the potential progression of adverse health  
9 effect in this community would benefit from additional research to establish the clinical  
10 significance of these findings. The observation by Whitehouse et al. (2008) of cases of  
11 mesothelioma among individuals with no direct occupational exposure to the mining and milling  
12 operations indicates the need for continued surveillance for this rare cancer.

#### 14 4.1.3. Marysville, OH Vermiculite Processing Plant Worker Studies

15 Libby vermiculite was used in the production of numerous commercial products,  
16 including as a potting soil amender and a carrier for pesticides and herbicides. A Marysville, OH  
17 plant that used Libby vermiculite in the production of fertilizer beginning around 1960 to 1980 is  
18 the location of the two related studies described in this section.

19 The processing facility had eight main departments, employing approximately  
20 530 workers, with 232 employed in production and packaging of the fertilizer and 99 in  
21 maintenance; other divisions included research, the front office, and the polyform plant (Lockey,  
22 1985). Six departments were located at the main facility (trionizing, packaging, warehouse,  
23 plant maintenance, central maintenance, and front offices). Research and development and a  
24 polyform fertilizer plant were located separately, approximately one-quarter mile from the main  
25 facility. In the trionizing section of the plant, the vermiculite ore was received by rail or truck,  
26 unloaded into a hopper, and transported to the expansion furnaces. After expansion, the  
27 vermiculite was blended with other materials (e.g., urea, potash, herbicides), packaged, and  
28 stored. Changes to the expander type and dust-control measures began in 1967, with substantial  
29 improvement in dust control occurring throughout the 1970s.

30 Information about exposure assessment at the Marysville, OH plant is summarized in the  
31 final row of Table 4-1. Industrial hygiene monitoring at the plant began in 1972. Lockey et al.

*This document is a draft for review purposes only and does not constitute Agency policy.*

Weinberg, 2011). Although limited, the data described in Section 4.2 suggest an increase in inflammatory response following exposure to Libby Amphibole asbestos and tremolite asbestos similar to that observed for other durable mineral fibers [reviewed in Mossman et al. (2007)]. Whether this inflammatory response then leads to cancer is unknown. Studies examining other types of asbestos (e.g., crocidolite, chrysotile, and amosite) have demonstrated an increase in chronic inflammation as well as respiratory cancer related to exposure [reviewed in Kamp and Weitzman (1999)]. Chronic inflammation has also been linked to genotoxicity and mutagenicity following exposure to some particles and fibers (Driscoll et al., 1997; 1996; 1995). The evidence described above suggests chronic inflammation is observed following Libby Amphibole asbestos and tremolite asbestos exposure; however, the role of inflammation and whether it leads to lung cancer or mesothelioma following exposure to Libby Amphibole asbestos is unknown.

ROS production has been measured in response to both Libby Amphibole asbestos and tremolite asbestos exposure. Blake et al. (2007) demonstrated an increase in the production of superoxide anion following exposure to Libby Amphibole asbestos. Blake et al. (2007) also demonstrated that total superoxide dismutase was inhibited, along with a decrease in intracellular glutathione, both of which are associated with increased levels of ROS. These results are supported by a recent study in human mesothelial cells (Hillegass et al., 2010) (described in Section 4.4 and Appendix D). Increased ROS production was also observed in human airway epithelial cells following exposure to Libby Amphibole asbestos (Duncan et al., 2010) (described in Section 4.4 and Appendix D). This increase in ROS and decrease in glutathione are common effects following exposure to asbestos fibers and particulate matter. Although ROS production is relevant to humans, based on similar human responses as compared to animals, information on the specifics of ROS production following exposure to Libby Amphibole asbestos is limited to the available data described here. Therefore, the role of ROS production in lung cancer and mesothelioma following exposure to Libby Amphibole asbestos is unknown.

#### 4.3. OTHER DURATION OR ENDPOINT-SPECIFIC STUDIES

##### 4.3.1. Immunological

Two epidemiology studies have examined the potential role of Libby Amphibole asbestos and autoimmunity. Noonan et al. (2006) used the data from the community health screening to examine self-reported history of autoimmune diseases (rheumatoid arthritis, scleroderma, or

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 lupus) in relation to the asbestos exposure pathways described above (see Table 4-17). To  
2 provide more specificity in the self-reported history of these diseases, a follow-up questionnaire  
3 was mailed to participants to confirm the initial report and obtain clarifying information  
4 regarding the type of disease, whether the condition had been diagnosed by a physician, and  
5 whether the participant was currently taking medication for the disease. Responses were  
6 obtained from 208 (42%) of the 494 individuals who had reported these conditions. Of these  
7 208 responses, 129 repeated the initial report of the diagnosis of rheumatoid arthritis, and  
8 161 repeated the initial report of the diagnosis of one of the three diseases (rheumatoid arthritis,  
9 scleroderma, or lupus). Among people aged 65 and over ( $n = 34$  rheumatoid arthritis cases,  
10 determined using responses from the follow-up questionnaire), a two- to threefold increase in  
11 risk was observed in association with several measures reflecting potential exposure to asbestos  
12 (e.g., asbestos exposure in the military) or specifically to Libby Amphibole asbestos (e.g., past  
13 work in mining and milling operations, use of vermiculite in gardening, and frequent playing on  
14 vermiculite piles when young). Restricted forced vital capacity, presence of parenchymal  
15 abnormalities, playing on vermiculite piles, and other dust or vermiculite exposures were also  
16 associated with rheumatoid arthritis in the group younger than 65 ( $n = 95$  cases). Restricted  
17 forced vital capacity was defined as FVC  $< 80\%$  predicted and a ratio of FEV1 to  
18 FVC  $\geq 70\%$  predicted. For all participants, an increased risk of rheumatoid arthritis was observed  
19 with increasing number of exposure pathways. RRs of 1.0, 1.02, 1.79, 2.51, and 3.98 were  
20 observed for 0 (referent), 1, 2–3, 4–5, and 6 or more pathways, respectively (trend  $p < 0.001$ ,  
21 adjusting for restrictive spirometry, parenchymal abnormalities, and smoking history). Although  
22 the information gathered in the follow-up questionnaire and repeated reports of certain diagnoses  
23 decreased the false-positive reports of disease, considerable misclassification (over-reporting and  
24 under-reporting) is likely, given the relatively low confirmation rate of self-reports of  
25 physician-diagnosed rheumatoid arthritis (and other autoimmune diseases) seen in other studies  
26 (Karlson et al., 2003; Rasch et al., 2003; Ling et al., 2000).

27 Another study examined serological measures of autoantibodies in 50 residents of Libby,  
28 MT, and a comparison group of residents of Missoula, Montana (Pfau et al., 2005); (see  
29 Table 4-17). The Libby residents were recruited for a study of genetic susceptibility to

30

Table 4-17. Autoimmune-related studies in the Libby, MT community

Reference(s)	Inclusion criteria and design details	Results
Noonan et al. (2006)	Nested case-control study among 7,307 participants in 2000–2001 community health screening. Conducted interviews, gathered self-reported history of rheumatoid arthritis, scleroderma, or lupus. Follow-up questionnaire mailed to participants concerning self-report of “physician-diagnosis” of these diseases and medication use.	Association with work in Libby mining/milling operations (ages 65 and older): Rheumatoid arthritis OR: 3.2 (95% CI: 1.3, 8.0) Rheumatoid arthritis, lupus, scleroderma OR: 2.1 (95% CI: 0.90, 4.1) Risk increased with increasing number of asbestos exposure pathways.
Pfau et al. (2005)	Libby residents ( $n = 50$ ) recruited for study of genetic susceptibility to asbestos-related lung disease. Missoula, MT comparison group ( $n = 50$ ), recruited for study of immune function; age and sex-matched to Libby participants. Serum samples obtained; IgA levels, prevalence of antinuclear, anti-dsDNA antibodies, anti-RF antibodies, and anti-Sm, RNP, SS-A, SS-B, and Scl-70 antibodies determined.	Increased prevalence of high titer ( $\geq 1:320$ ) antinuclear antibodies in Libby sample (22%) compared to Missoula sample (6%). Similar increases for rheumatoid factor, anti-RNP, anti-Scl-60, anti-Sm, anti-R <sub>o</sub> (SSA), and anti-La (SSB) antibodies observed in Libby sample.

asbestos-related lung disease, and the Missoula residents were participants in a study of immune function. The Libby sample exhibited an increased prevalence (22%) of high-titer ( $\geq 1:320$ ) antinuclear antibodies when compared to the Missoula sample (6%), and similar increases were seen in the Libby sample for rheumatoid factor, anti-RNP, anti-Scl-60, anti-Sm, anti-R<sub>o</sub> (SSA), and anti-La (SSB) antibodies. Although neither sample was randomly selected from the community residents, an individual’s interest in participating in a gene and lung disease study likely would not be influenced by the presence of autoimmune disease or autoantibodies in that individual.

Hamilton et al. (2004), Blake et al. (2008), and Pfau et al. (2008) examined the role of asbestos in autoimmunity in laboratory animal or in vitro studies. Blake et al. (2008) performed in vitro assays with Libby Amphibole asbestos (see Section 4.4), and both studies performed the in vivo assays with tremolite. C57BL/6 mice were instilled intratracheally for a total of two doses each of 60- $\mu$ g saline and wollastonite or Korean tremolite sonicated in sterile PBS, given 1 week apart in the first 2 weeks of a 7-month experiment. Sera from mice exposed to tremolite showed antibody binding colocalized with SSA/Ro52 on the surface of apoptotic blebs (Blake et al., 2008). In Pfau et al. (2008), by 26 weeks, the tremolite-exposed animals had a significantly

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 Tremolite and Libby Amphibole asbestos exposure led to increases in both fibrosis and  
2 tumorigenicity in all but one animal study, supporting a possible role for proliferation in  
3 response to these fibers. However, there are limited data to demonstrate that increased  
4 cytotoxicity and cellular proliferation following exposure to Libby Amphibole asbestos leads to  
5 lung cancer or mesothelioma.

6 *Summary.* The review of these studies clearly highlights the need for more controlled  
7 studies examining Libby Amphibole asbestos in comparison with other forms of asbestos and for  
8 examining multiple endpoints—including ROS production, DNA damage, and pro-inflammatory  
9 gene expression alterations—to improve understanding of mechanisms involved in cancer and  
10 other health effects. Data gaps still remain to determine specific mechanisms involved in Libby  
11 Amphibole asbestos-induced disease. Studies that examined cellular response to tremolite also  
12 found that tremolite exposure may lead to increased ROS production, toxicity, and genotoxicity  
13 (Okayasu et al., 1999; Wagner et al., 1982). As with the in vivo studies, the definition of fibers  
14 and how the exposures were measured varies among studies.

#### 16 4.5. SYNTHESIS OF MAJOR NONCANCER EFFECTS

17 The predominant noncancer health effects observed following inhalation exposure to  
18 Libby Amphibole asbestos are effects on the lungs and pleural lining surrounding the lungs.  
19 Recent studies have also examined noncancer health effects following exposure to Libby  
20 Amphibole asbestos in other systems, including autoimmune effects and cardiovascular disease.  
21 These effects have been observed primarily in studies of exposed workers and community  
22 members and are supported by laboratory animal studies.

##### 24 4.5.1. Pulmonary Effects

###### 25 4.5.1.1. Pulmonary Fibrosis (Asbestosis)

26 Asbestosis is the interstitial pneumonitis and fibrosis caused by inhalation of asbestos  
27 fibers and is characterized by a diffuse increase of collagen in the alveolar walls (fibrosis) and  
28 the presence of asbestos fibers, either free or coated with a proteinaceous material and iron  
29 (asbestos bodies). Fibrosis results from a sequence of events following lung injury, which  
30 includes inflammatory cell migration, edema, cellular proliferation, and accumulation of  
31 collagen. Asbestosis is associated with dyspnea, bibasilar rales, and changes in pulmonary

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 function: a restrictive pattern, mixed restrictive-obstructive pattern, and/or decreased diffusing  
2 capacity (ATS, 2004). Radiographic evidence of small opacities in the lung is direct evidence of  
3 scarring of the lung tissue and as the fibrotic scarring of lung tissue consistent with mineral dust  
4 and mineral fiber toxicity. The scarring of the parenchymal tissue of the lung contributes to  
5 measured changes in pulmonary function, including obstructive pulmonary deficits from  
6 narrowing airways, restrictive pulmonary deficits from impacting the elasticity of the lung as  
7 well as decrements in gas exchange.

8 Workers exposed to Libby Amphibole asbestos from vermiculite mining and processing  
9 facilities in Libby, MT, as well as plant workers in Marysville, OH, where vermiculite ore was  
10 exfoliated and processed, have an increased prevalence of small opacities on chest X-rays, which  
11 is indicative of fibrotic damage to the parenchymal tissue of the lung (Rohs et al., 2008;  
12 Amandus et al., 1987b; McDonald et al., 1986b; Lockey et al., 1984). These findings are  
13 consistent with a diagnosis of asbestosis, and the studies are described in detail in  
14 Section 4.1.1.4.2. Significant increases in asbestosis as the primary cause-of-death have been  
15 documented in studies of the Libby worker cohort report (see Table 4.6 for details) (Larson et al.,  
16 2010b; Sullivan, 2007; Amandus and Wheeler, 1987; McDonald et al., 1986a). For both  
17 asbestosis mortality and radiographic signs of asbestos (small opacities), positive exposure-  
18 response relationships are described where these effects are greater with greater cumulative  
19 exposure to Libby Amphibole asbestos.

20 Deficits in pulmonary function consistent with pulmonary fibrosis have been reported in  
21 individuals exposed to Libby Amphibole asbestos. The initial study of the Marysville, OH  
22 cohort measured but reported no change in pulmonary function (Lockey et al., 1984).  
23 Pulmonary function was not reported for the cohort follow-up, although prevalence of pleural  
24 and parenchymal abnormalities was increased (Rohs et al., 2008). Although studies of the  
25 occupational Libby worker cohort do not include assessment of pulmonary function (Amandus et  
26 al., 1987b; McDonald et al., 1986b) data from the ATSDR community screening, which included  
27 workers, provide support for functional effects from parenchymal changes. The original report  
28 of the health screening data indicated moderate-to-severe pulmonary restriction in 2.2% of men  
29 (Peipins et al., 2003; ATSDR, 2001b). A recent reanalysis of these data show that for study  
30 participants with small opacities viewed on the radiographs (grade 1/0 or greater), and DPT the  
31 mean FVC is reduced to 78.76 ( $\pm 3.64$ ), 82.16 ( $\pm 3.34$ ), respectively of the expected value (Weill

*This document is a draft for review purposes only and does not constitute Agency policy.*

et al., 2011). A mean FVC of 95.63 ( $\pm 0.76$ ) was reported for those with other pleural abnormalities versus 103.15 ( $\pm 0.25$ ) in participants with no radiographic abnormalities. The strongest effects of diffuse pleural thickening and/or costophrenic angle obliteration on FVC were seen among men who had never smoked ( $-23.77$ ,  $p < 0.05$ ), with smaller effects seen among men who had smoked ( $-9.77$ ,  $p < 0.05$ ) and women who had smoked ( $-6.73$ ,  $p < 0.05$ ). Laboratory animal and mechanistic studies of Libby Amphibole asbestos are consistent with the noncancer health effects observed in both Libby workers and community members. Pleural fibrosis was increased in hamsters after intrapleural injections of Libby Amphibole asbestos (Smith, 1978). More recent studies have demonstrated increased collagen deposition consistent with fibrosis following intratracheal instillation of Libby Amphibole asbestos fibers in mice (Padilla-Carlin et al., 2011; Shannahan et al., 2011a; Shannahan et al., 2011b; Smartt et al., 2010; Putnam et al., 2008). Pulmonary fibrosis, inflammation, and granulomas were observed after tremolite inhalation exposure in Wistar rats (Bernstein et al., 2005; Bernstein et al., 2003) and intratracheal instillation in albino Swiss mice (Sahu et al., 1975). Davis et al. (1985) also reported pulmonary effects after inhalation exposure in Wistar rats including increases in peribronchiolar fibrosis, alveolar wall thickening, and interstitial fibrosis.

#### 4.5.1.2. Other Nonmalignant Respiratory Diseases

Mortality studies of the Libby workers indicate that there is increased mortality, not only from asbestosis, but other respiratory diseases. Deaths attributed to chronic obstructive respiratory disease and deaths attributed to “other” nonmalignant respiratory disease were elevated more than twofold (see Table 4-6) (Larson et al., 2010b; Sullivan, 2007). These diseases are consistent with asbestos toxicity, and the evidence of a positive exposure-response relationship for mortality from all nonmalignant respiratory diseases, supports this association.

#### 4.5.2. Pleural Effects

Pleural thickening that is caused by mineral fiber exposure includes two distinct biological lesions: discrete pleural plaques in the parietal pleura and diffuse pleural thickening of the visceral pleura. Both forms of pleural thickening can be viewed on standard radiographs. However, the two are not always clearly distinguishable on X-rays, and smaller lesions may not be detected. High resolution computed tomography is a method that can distinguish between the

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 lesions, as well as detect smaller lesions than are visible on X-rays. Pleural thickening may  
2 restrict lung function, increase breathlessness with exercise, and contribute to chronic chest pain.  
3 The potential for health effects and severity of health effects are increased with the extent and  
4 thickness of the pleural lesions.

5 Data from the ATSDR community health screening study indicate that the prevalence of  
6 pleural abnormalities, identified by radiographic examination, increases substantially with  
7 increasing number of exposure pathways (Peipins et al., 2003). A reanalysis of these data also  
8 considered age, smoking history, and types of exposures. Increased pleural thickening is  
9 reported for Libby workers, those with other vermiculite work and those in “dusty trades.”  
10 Increased LPT is reported in both those exposed only as household contacts or through  
11 environmental exposure pathways, with greater incidence by age (38.3 and 12.7%, respectively,  
12 in the 61–90 age group) (Weill et al., 2011). DPT is reported at lower rates with 5.9 and 2.2%,  
13 respectively, in these exposure groups in the highest age bracket evaluated (age 61–90).

14 Increased pleural thickening is reported for both of the studied worker cohorts, with  
15 evidence of positive exposure response relationships (Larson et al., 2010a; Rohs et al., 2008;  
16 Amandus et al., 1987b; McDonald et al., 1986b; Lockey et al., 1984). Both McDonald et al.  
17 (1986b) and Amandus et al. (1987b) indicate age is also a predictor of pleural thickening in  
18 exposed individuals, which may reflect the effects of time from first exposure. Smoking data  
19 were limited on the Libby workers and analyses do not indicate clear relationships between  
20 smoking and pleural thickening (Amandus et al., 1987b; McDonald et al., 1986b). Pleural  
21 thickening in workers at the Scott Plant (Marysville, OH) was associated with hire on or before  
22 1973 and age at time of interview but was not associated with BMI or smoking history (ever  
23 smoked) (Rohs et al., 2008).

#### 24 25 **4.5.3. Other Noncancer Health Effects (Cardiovascular Toxicity, Autoimmune Effects)**

26 There is limited research available on noncancer health effects occurring outside the  
27 respiratory system. Larson et al. (2010b) examined cardiovascular disease-related mortality in  
28 the cohort of exposed workers from Libby (see Section 4.1.1.4.3). Mechanistic studies have  
29 examined the potential role of iron and the associated inflammation for both the respiratory and  
30 cardiovascular disease (Shannahan et al., 2011b). Two studies examined the association between  
31 asbestos exposure and autoimmune disease (Noonan et al., 2006) or autoantibodies and other

*This document is a draft for review purposes only and does not constitute Agency policy.*

immune markers (Pfau et al., 2005) (see Table 4-17). Limitations in the number, scope, and design of these studies make it difficult to reach conclusions as to the role of asbestos exposure in either cardiovascular disease or autoimmune disease.

#### 4.5.4. Libby Amphibole Asbestos Summary of Noncancer Health Effects

The studies in humans summarized in Section 4.1 have documented an increase in mortality from nonmalignant respiratory disease, including asbestosis, in workers exposed to Libby Amphibole asbestos (Larson et al., 2010b; Sullivan, 2007; McDonald et al., 2004; Amandus and Wheeler, 1987). Radiographic evidence of pleural thickening and interstitial damage (small opacities) are also well documented among employees of the Libby vermiculite mining operations (Larson et al., 2010a; Amandus et al., 1987b; McDonald et al., 1986b). Additional studies have documented an increase in radiographic changes in the pleura and parenchyma among employees of a manufacturing facility in Marysville, OH that used Libby vermiculite ore contaminated with Libby Amphibole asbestos (Rohs et al., 2008; Lockey et al., 1984). Positive exposure-response relationships for these health effects for both occupational cohorts studied, as well as the observed latency, support an association between exposure to Libby Amphibole asbestos and these pleuro-pulmonary effects. Studies of community members exposed to Libby Amphibole asbestos have documented similar pleural abnormalities and pulmonary deficits consistent with parenchymal damage (Weill et al., 2011; Whitehouse, 2004; Peipins et al., 2003). Although limited, animal studies support the toxicity of Libby Amphibole asbestos to pleural and pulmonary tissues. Developing research supports a role of inflammatory processes in the toxic action of Libby Amphibole asbestos, consistent with the observed health effects (Duncan et al., 2010; Hamilton et al., 2004). Taken together, the strong evidence in human studies, defined exposure response relationships, and supportive animal studies provide compelling evidence that exposure to Libby Amphibole asbestos causes nonmalignant respiratory disease, including asbestosis, pleural thickening, and deficits in pulmonary function associated with mineral fiber exposures. Existing data regarding cardiovascular effects and the potential for autoimmune disease are limited.

#### 4.5.5. Mode-of-Action Information (Noncancer)

The precise mechanisms causing toxic injury from inhalation exposure to Libby Amphibole asbestos have not been established. However, nearly all-durable mineral fibers with dimensional characteristics that allow penetration to the terminal bronchioles and alveoli of the lung have the capacity to induce pathologic response in the lung and pleural cavity (ATSDR, 2001a; Witschi and Last, 1996). The physical-chemical attributes of mineral fibers are important in determining the type of toxicity observed. Fiber dimension (width and length), density, and other characteristics such as chemical composition, surface area, solubility in physiological fluids, and durability all play important roles in both the type of toxicity observed and the biologically significant dose. Fibrosis results from a sequence of events following lung injury, which includes inflammatory cell migration, edema, cellular proliferation, and accumulation of collagen. Fibers do migrate to the pleural space, and it has been hypothesized that a similar cascade of inflammatory events may contribute to fibrotic lesions in the visceral pleura. Thickening of the visceral pleura is more often localized to lobes of the lung with pronounced parenchymal changes, and it has also been hypothesized that the inflammatory and fibrogenic processes within the lung parenchyma in response to asbestos fibers may influence the fibrogenic process in the visceral pleura. The etiology of parietal plaques is largely unknown with respect to mineral fiber exposure.

There is currently insufficient evidence to establish the noncancer mode of action for Libby Amphibole asbestos. Limited in vitro studies have demonstrated oxidative stress following Libby Amphibole asbestos exposures in various cell types (Duncan et al., 2010; Hillegass et al., 2010; Pietruska et al., 2010; Blake et al., 2007). Libby Amphibole asbestos fibers increased intracellular ROS in both murine macrophages and human epithelial cells (Duncan et al., 2010; Blake et al., 2007). Surface iron, inflammatory marker gene expression was increased following exposure to Libby Amphibole asbestos in human epithelial cells (Shannahan et al., 2011b; Duncan et al., 2010; Pietruska et al., 2010) (see Table 4-18). Tremolite studies demonstrate cytotoxicity in various cell culture systems (see Table 4-19).

The initial stages of any fibrotic response involve cellular proliferation, which may be compensatory for cell death due to cytotoxicity. Analysis of cellular proliferation has demonstrated both increases and decreases following exposure to asbestos fibers in vitro and in vivo depending on the specific fiber or cell type (Mossman et al., 1985; Topping and Nettesheim,

*This document is a draft for review purposes only and does not constitute Agency policy.*

180). Other studies have focused on the activation of cell-signaling pathways that lead to cellular proliferation following exposure to asbestos (Scapoli et al., 2004; Shukla et al., 2003; Ding et al., 1999; Zanella et al., 1996).

Although slightly increased compared to controls, cytotoxicity in murine macrophage cells exposed to Libby Amphibole asbestos was decreased compared to other fiber types (Blake et al., 2008). Cytotoxicity was slightly, but statistically significantly, increased compared to an unexposed control at 24 hours post exposure to Libby Amphibole asbestos, while crocidolite exposure resulted in even higher levels of cytotoxicity. No other in vitro study examined cytotoxicity following exposure to Libby Amphibole asbestos, although an increase in apoptosis was demonstrated in this same cell system (Blake et al., 2008). Recent studies in mice exposed to Libby Amphibole asbestos demonstrated increased collagen deposition and collagen gene expression, markers of fibrosis (Smartt et al., 2010; Putnam et al., 2008). Short-term studies in rats also demonstrated an increased inflammatory response (Padilla-Carlin et al., 2011; Shannahan et al., 2011a; Shannahan et al., 2011b). Tremolite and Libby Amphibole asbestos exposure led to increases in both fibrosis in all but one animal study, supporting a role for proliferation in response to these fibers. Taken together with studies on other asbestos fibers, these data suggest that a cytotoxicity and cell proliferation may play a role in the noncancer health effects following exposure to Libby Amphibole asbestos.

Although continued research demonstrates that the Libby Amphibole asbestos has biologic activity consistent with the inflammatory action and cytotoxic effects seen with other forms of asbestos, the data are not sufficient to establish a mode of action for the pleura-pulmonary effects of exposure to Libby Amphibole asbestos.

## 4.6. EVALUATION OF CARCINOGENICITY

### 4.6.1. Summary of Overall Weight of Evidence

Under the EPA *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005a), Libby Amphibole asbestos is *carcinogenic to humans* following inhalation exposure based on epidemiologic evidence that shows a convincing association between exposure to Libby Amphibole asbestos fibers and increased lung cancer and mesothelioma mortality (Larson et al., 2010b; Moolgavkar et al., 2010; Sullivan, 2007; McDonald et al., 2004; Amandus and Wheeler, 1987; McDonald et al., 1986a). These results are further supported by animal studies that

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 Table 5-2. Summary of rationale for identifying candidate principal studies  
 2 on Libby Amphibole asbestos for RfC development  
 3

Attribute	Preferred characteristics for candidate principal studies for the Libby Amphibole Asbestos RfC
Relevance of exposure paradigm	<p>Studies of subchronic or chronic duration are preferred over studies of acute exposure duration because most relevant environmental exposure scenarios are expected to address chronic exposure scenarios (potentially including both continuous exposure from ambient conditions and episodic activity-related exposures).</p> <p>Measures of cumulative exposure are a widely used metric to address asbestos risk. It is consistent with the expectation that toxic responses will reflect an accumulative effect of asbestos inhaled and deposited in tissues over time. Additionally mean exposure, exposure duration, and time from first exposure (TSFE) have all been reported as predictors of health effects from asbestos exposure. Cumulative exposure has the advantage that it reflects both duration and intensity (e.g., mean level) of asbestos exposure.</p> <p>Relatively lower exposure intensities that may represent conditions more similar to environmental exposures are preferred as there may be less uncertainty in extrapolation of the results to lower exposure levels.</p> <p>Results from studies with high exposure intensity or cumulative exposure are, other things being comparable, judged less relevant for environmental risk assessment compared to studies defining effects at lower levels of exposure. Some biological processes (e.g., potential decrease in effectiveness of particle clearance processes) may more strongly influence responses at very high levels of exposure and be less relevant at lower levels. Thus, exposure conditions with lower level exposures may remove some of the uncertainty in estimating health effects from environmental exposures.</p>
Study design characteristics	<p>Sufficient follow-up time for outcomes to develop (which can depend on the health outcome being addressed).</p> <p>Study size and participation rates that are adequate to detect and quantify health outcomes being studied are preferred, with no indications of bias in study population selection.</p> <p>Use of a study design or analytic approach, which adequately addresses the relevant sources of potential confounding, including age, sex, smoking, and exposure to other risk factors (such as non-Libby asbestos).</p>

4

*This document is a draft for review purposes only and does not constitute Agency policy.*

**Table 5-2. Summary of rationale for identifying candidate principal studies on Libby Amphibole asbestos for RfC development (continued)**

Measurement of exposure	<p>Emphasis is placed on the specificity of exposure assessment in time and place with a preference for greater detail where possible. Exposure measurements that are site- and task-specific provide appropriate exposure information, and individual, rather than area samples are preferred where available. Measurement techniques that are more specific to the agent of concern are preferred over less specific analytical methods. Better characterization of fibers is preferred. For asbestos fibers, TEM analysis, which can identify the mineral fibers present, provides the most specific information; PCM identifies fibers as defined by that method (NIOSH 7400) and, thus, is useful but do not confirm the mineral nature of the counted fibers. Total dust measurements are the least informative of those available.</p> <p>Stronger studies will often be based upon knowledge of individual work histories (job titles/tasks with consideration of changes over time); however, appropriate group-based exposure estimates may also be relevant.</p> <p>Exposure reconstruction and estimating exposures based on air sampling from other time periods and/or operations are less preferred methods of exposure estimation.</p>
Measurement of effect(s)	<p>Emphasis is placed on the more sensitive health outcome endpoints that are available. For parenchymal and pleural effects considered here, the radiographic abnormalities are more sensitive than the corresponding mortality causes. An RfC is intended to be a level at which no category of adverse health outcome would occur.</p> <p>Pleural and parenchymal abnormalities assessed using good quality radiographs or high-resolution computed tomography (HRCT) and independently evaluated multiple qualified readers according to ILO standards.</p> <p>Evaluation of radiographs should not be influenced by knowledge of exposure status.</p>

intensity exposures for the Marysville cohort and corresponding lower cumulative exposures are advantages of this study, considering there are uncertainties inherent in exposure-response data and extrapolating from the high intensity occupation exposures to lower level exposures often seen in community and environmental exposures.

#### **5.2.1.2.1. Evaluation of study design in candidate studies**

The candidate principal studies differed in the study populations, in terms of follow-up time, study size and participation, and available information (see Table 5-1). The study sizes are similar for the two Libby worker studies ( $n = 184$  and  $n = 244$ , respectively) (Amandus et al., 1987b; McDonald et al., 1986b) and the Marysville update ( $n = 280$ ) (Rohs et al., 2008).

Adequate follow-up time allows for the health effect to manifest prior to sampling. In the case of pleural abnormalities, there is some variability with latency based on intensity of

*This document is a draft for review purposes only and does not constitute Agency policy.*

1 exposure as well as the nature of the pleural lesion where discrete pleural plaques have a shorter  
2 latency than diffuse thickening of the visceral pleura. Larson et al. (2010a) studied the latency  
3 for individuals in the Libby worker cohort, reporting a median latency of 8.6 years for localized  
4 pleural thickening versus 27 years for diffuse pleural thickening and 19 years for minimal signs  
5 of small opacities (parenchymal changes).<sup>24</sup> Lockey et al. (1984) report the mean employment  
6 duration for their exposure groups from 6.6 to 13.3 years at the time of their study (but do not  
7 assess time since first exposure (TSFE); thus, it is unclear whether in the first examination these  
8 workers had sufficient follow-up to assess the radiographic changes, especially diffuse pleural  
9 thickening and small opacities. The Rohs et al. (2008) report includes 24 more years of  
10 follow-up time and is preferred over the early Lockey et al. (1984) study on this basis.

11 Both studies of the Libby workers report duration of employment and average age of the  
12 participants, but not TSFE. The McDonald et al. (1986b) study included both current and former  
13 workers—these former workers likely have longer time from first exposure compared with  
14 current workers. The study included all current plant employees (164 men, 9 women).  
15 However, there was a lower participation rate in former employees (80 of 110 eligible former  
16 employees agreed to provide chest radiographs). Additionally, X-rays for all study participants  
17 were taken in the same year, providing similar quality X-rays between past and current  
18 employees. In contrast, Amandus et al. (1987b) only considered workers employed during 1975  
19 to 1982 and relied on available radiographs regardless of year (radiographs were available for  
20 93% of employees). Because workers terminated prior to 1975 were excluded from the study,  
21 older individuals, and individuals with longer TSFE were less likely to be included than in the  
22 study by McDonald et al. (1986b), which included former workers. Both Libby worker studies  
23 do report radiographic abnormalities, so the follow-up is adequate for some effects to be  
24 documented; however, compared with the Rohs et al. (2008) study, the Libby worker studies  
25 have shorter follow-up times.

---

<sup>24</sup> Individual latency for visible LPT in Libby exposed workers was evaluated in 84 workers with radiographic evidence of pleural and/or parenchymal changes (Larson et al., 2010a). By examining historical radiographs, researchers were able to identify the first appearance of the lesions, although it is recognized that retrospective design of this study likely identified lesions at earlier time points, as the readers were aware of the later X-rays (Larson et al., 2010a). It is acknowledged that some of the workers at Libby may have been exposed through the community prior to working, and in fact, one individual had the first pleural change noted at 9 years of age, prior to occupational exposure (Larson et al., 2010a). Where data on prior exposures were available, workers with no prior exposure had an average latency of 9.4 years versus 5.1 years for workers with potential exposures prior to hire (N = 63 and 31, respectively).